BEST AVAILABLE COPY

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07D 403/14, 207/34, A61K 31/415, C12Q 1/68

A1

(11) International Publication Number: WO 98/45284

(43) International Publication Date: 15 October 1998 (15.10.98)

(21) International Application Number: PCT/US98/03829

(22) International Filing Date: 29 January 1998 (29.01.98)

(30) Priority Data:

| | | |
|---|---|---|
| PCT/US97/03332 | 20 February 1997 (20.02.97) | US |
| 60/043,444 | 8 April 1997 (08.04.97) | US |
| 60/042,022 | 16 April 1997 (16.04.97) | US |
| 08/837,524 | 21 April 1997 (21.04.97) | US |
| 08/853,522 | 8 May 1997 (08.05.97) | US |
| PCT/US97/12722 | 21 July 1997 (21.07.97) | US |

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

| | |
|---|---|
| US | 08/853,522 (CIP) |
| Filed on | 8 May 1997 (08.05.97) |
| US | 08/837,524 (CIP) |
| Filed on | 21 April 1997 (21.04.97) |
| US | 08/607,078 (CIP) |
| Filed on | 26 February 1996 (26.02.96) |
| US | 60/042,022 (CIP) |
| Filed on | 16 April 1997 (16.04.97) |
| US | 60/043,444 (CIP) |
| Filed on | 8 April 1997 (08.04.97) |

(71) Applicant *(for all designated States except US)*: CALIFORNIA INSTITUTE OF TECHNOLOGY [US/US]; Pasadena, CA 91125 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: BAIRD, Eldon, E. [US/US]; 255 S. Madison Avenue #5, Pasadena, CA 91101 (US). DERVAN, Peter, B. [US/US]; 1235 St. Albans Road, San Marino, CA 91108 (US).

(74) Agent: McDONNELL, John, J.; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

**Published**

*With international search report.*

*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: STEREOCHEMICAL CONTROL OF THE DNA BINDING AFFINITY, SEQUENCE SPECIFICITY, AND ORIENTATION–PREFERENCE OF CHIRAL HAIRPIN POLYAMIDES IN THE MINOR GROOVE

(57) Abstract

This invention provides improved polyamides comprising a hairpin loop derived from γ–aminobutyric acid which bind to the minor groove of a promoter regions of a DNA sequence. Binding of the polyamide to the DNA sequence of the promoter region inhibits expression of the requisite gene. The improvement relates to the use of R–2,4–diaminobutyric acid and derivatives of the 2–amino group to form the hairpin loop. The improved asymmetric hairpin provides for tighter binding of the polyamides to the minor groove of DNA and additionally provides an amine function for derivatizing polyamides by, for example, forming amide linkages. Such derivatives may serve to attach detectable labels to the polyamide.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **AL** | Albania | **ES** | Spain | **LS** | Lesotho | **SI** | Slovenia |
| **AM** | Armenia | **FI** | Finland | **LT** | Lithuania | **SK** | Slovakia |
| **AT** | Austria | **FR** | France | **LU** | Luxembourg | **SN** | Senegal |
| **AU** | Australia | **GA** | Gabon | **LV** | Latvia | **SZ** | Swaziland |
| **AZ** | Azerbaijan | **GB** | United Kingdom | **MC** | Monaco | **TD** | Chad |
| **BA** | Bosnia and Herzegovina | **GE** | Georgia | **MD** | Republic of Moldova | **TG** | Togo |
| **BB** | Barbados | **GH** | Ghana | **MG** | Madagascar | **TJ** | Tajikistan |
| **BE** | Belgium | **GN** | Guinea | **MK** | The former Yugoslav Republic of Macedonia | **TM** | Turkmenistan |
| **BF** | Burkina Faso | **GR** | Greece | | | **TR** | Turkey |
| **BG** | Bulgaria | **HU** | Hungary | **ML** | Mali | **TT** | Trinidad and Tobago |
| **BJ** | Benin | **IE** | Ireland | **MN** | Mongolia | **UA** | Ukraine |
| **BR** | Brazil | **IL** | Israel | **MR** | Mauritania | **UG** | Uganda |
| **BY** | Belarus | **IS** | Iceland | **MW** | Malawi | **US** | United States of America |
| **CA** | Canada | **IT** | Italy | **MX** | Mexico | **UZ** | Uzbekistan |
| **CF** | Central African Republic | **JP** | Japan | **NE** | Niger | **VN** | Viet Nam |
| **CG** | Congo | **KE** | Kenya | **NL** | Netherlands | **YU** | Yugoslavia |
| **CH** | Switzerland | **KG** | Kyrgyzstan | **NO** | Norway | **ZW** | Zimbabwe |
| **CI** | Côte d'Ivoire | **KP** | Democratic People's Republic of Korea | **NZ** | New Zealand | | |
| **CM** | Cameroon | | | **PL** | Poland | | |
| **CN** | China | **KR** | Republic of Korea | **PT** | Portugal | | |
| **CU** | Cuba | **KZ** | Kazakstan | **RO** | Romania | | |
| **CZ** | Czech Republic | **LC** | Saint Lucia | **RU** | Russian Federation | | |
| **DE** | Germany | **LI** | Liechtenstein | **SD** | Sudan | | |
| **DK** | Denmark | **LK** | Sri Lanka | **SE** | Sweden | | |
| **EE** | Estonia | **LR** | Liberia | **SG** | Singapore | | |

# Stereochemical Control of the DNA Binding Affinity, Sequence Specificity, and Orientation-Preference of Chiral Hairpin Polyamides in the Minor Groove

The U.S. Government has certain rights to this invention pursuant to Grant Nos. GM 26453, 27681, and 47530 awarded by the National Institute of Health.

## CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of PCT/US97/03332 filed February 20, 1997, Serial No. 08/853,522 filed May 8, 1997 and PCT/US97/12722 filed July 21, 1997 which are continuation-in-part applications of Serial No. 08/837,524 filed April 21, 1997 and Serial No. 08/607,078 filed February 26, 1996; and provisional application 60/042,022, filed April 16, 1997 and provisional application 60/043,444 filed April 8, 1997. The specification of these applications are incorporated herein by reference.

## BACKGROUND OF THE INVENTION

### Field of the Invention

This invention relates to polyamides which bind to pre-determined sites of the minor groove of double-stranded DNA.

### Description of the Related Art

The art describes a large variety of polyamides which have three to six carboxamide base pairs and a hairpin loop derived from γ-aminobutyric acid and the ability to bind to the minor groove of DNA in the promoter region to inhibit gene expression. Thus, polyamides consisting of N-methylimidazole (Im), N-methylpyrrole (Py), and β-alanine and γ-amino butyric acid and methods for preparation of such polyamides are well known.

Polyamides containing *N*-methylpyrrole and *N*-methylimidazole amino acids are synthetic ligands that have an affinity and specificity for DNA comparable to naturally

occurring DNA binding proteins (Trauger, et al. *Nature* **1996**, *382*, 559; Swalley, et al. *J. Am. Chem. Soc.* **1997**, *119*, 6953; Turner, et al. *J. Am. Chem. Soc.* **1997**, *119*, 7636). DNA recognition depends on side-by-side amino acid pairings oriented N-C with respect to the 5'-3' direction of the DNA helix in the minor groove (Wade, W. S., et al. *J. Am. Chem. Soc.* **1992**, *114*, 8783; Mrksich, et al. *Proc. Natl. Acad. Sci., USA* **1992**, *89*, 7586; Wade, et al. *Biochemistry* **1993**, *32*, 11385; Mrksich, et al. *J. Am. Chem. Soc.* **1993**, *115*, 2572; Geierstanger, et al. *Science* **1994**, *266*, 646; White, et al. *J. Am. Chem. Soc.* **1997**, *119*, 8756). Antiparallel pairing of imidazole (Im) opposite pyrrole (Py) recognizes a G•C base pair, while a Py/Im combination recognizes C•G.[2] A Py/Py pair is degenerate and recognizes either an A•T or T•A base pair (Wade, W. S., et al. *J. Am. Chem. Soc.* **1992**, *114*, 8783; Mrksich, et al. *Proc. Natl. Acad. Sci., USA* **1992**, *89*, 7586; Wade, et al. *Biochemistry* **1993**, *32*, 11385; Mrksich, et al. *J. Am. Chem. Soc.* **1993**, *115*, 2572; Geierstanger, et al. *Science* **1994**, *266*, 646; White, et al. *J. Am. Chem. Soc.* **1997**, *119*, 8756; Pelton, et al. *Proc. Natl. Acad. Sci., USA* **1989**, *86*, 5723; Pelton, et al. *J. Am. Chem. Soc.* **1990**, *112*, 1393; White, et al. *Biochemistry* **1996**, *35*, 12532; Chen, et al. *J. Mol. Biol.* **1997**, *267*, 1157). An Im/Im pairing is disfavored, breaking a potential degeneracy for recognition (Singh, et al. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7673; White, et al. *Chem. & Biol.* **1997**, *4*, 569).

Investigators have also attempted to prevent slipped-binding motifs as well as increase DNA-binding affinity and sequence specificity by covalent linkage of polyamide subunits (Trauger, et al. *J. Am. Chem. Soc.* **1996**, *118*, 6160; Geierstanger, et al. *Nature Struct. Biol.* **1996**, *3*, 321; Swalley, et al. *Chem. Eur. J.* **1997**, *3*, 1608; Wemmer, et al. *Curr . Opin. Struct. Biol.* **1997**, *7*, 355; Mrksich, et al. *J. Am. Chem. Soc.* **1994**, *116*, 3663; Dwyer, et al. *J. Am. Chem. Soc.* **1993**, *115*, 9900; Chen, et al. *J. Am. Chem. Soc.* **1994**, *116*, 6995). A hairpin polyamide motif with γ-aminobutyric acid (γ) has been utilized as a turn-specific internal-guide-residue and provides a synthetically accessible method for C-N linkage of polyamide subunits (Figure 1). Head-to-tail linked

polyamides bind specifically to designated target sites with 100-fold enhanced affinity relative to unlinked subunits (Mrksich, et al. *J. Am. Chem. Soc.* **1994**, *116*, 7983; Parks, et al. *J. Am. Chem. Soc.* **1996**, *118*, 6147; Parks, et al. *J. Am. Chem. Soc.* **1996**, *118*, 6153; Trauger, et al. *Chem. & Biol.* **1996**, *3*, 369; Swalley, et al. *J. Am. Chem. Soc.* **1996**, *118*, 8198; Pilch, et al. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306; de Claire, et al. *J. Am. Chem. Soc.* **1997**, *119*, 7909).

Eight-ring hairpin polyamides bearing a single positively charged tertiary amine group at the C-terminus have been shown to be cell-permeable and to inhibit the transcription of specific genes in cell culture (Gottesfeld, et al. *Nature* **1997**, *387*, 202). However, recent studies of polyamide size limitations suggest that beyond five rings, the ligand curvature fails to match the pitch of the DNA helix, disrupting the hydrogen bonds and van der Waals interactions responsible for specific polyamide-DNA complex formation (Kelley, et al. *Proc. Natl. Acad. Sci. USA*, **1996**, 93:6981; Kielkopf, et al. *Nature Struc. Biol.*, in press). Recognition of seven base pairs by ten-ring hairpin polyamids containing five contiguous ring pairings represents the upper limit in binding site sizes targetable by the hairpin motif (Turner, et al. *J. am. Chem. Soc.*, **1997**, 119:7636). Addition of pairings of β-alanine with β-alanine, pyrrole, or imidazole has allowed extention of the hairpin motif to 8-bp recognition, as demonstrated in provisional application 60/042,222. However, those skilled in the art have recognized the extreme difficulties associated with the design of hairpin motifs recognizing longer site sizes.

The present invention involves the use of R-2,4-diaminobutyric acid as a replacement for γ-aminobutyric acid to make the hairpin loop. In addition, a methodology for expanding the targetable binding site size of hairpins by covalently linking existing hairpin motifs without compromising DNA-binding and sequence specificity is provided.

## SUMMARY OF THE INVENTION

This invention provides improved polyamides comprising a hairpin loop derived from γ-aminobutyric acid which bind to the minor groove of a promoter region of a DNA sequence. Binding of the polyamide to the DNA sequence of the promoter region inhibits expression of the requisite gene. The improvement relates to the use of R-2,4-diaminobutyric acid and derivatives of the 2-amino group to form the hairpin loop. The improved asymmetric hairpin provides for tighter binding of the polyamides to the minor groove of DNA and additionally provides an amine function for derivatizing polyamides by, for example, forming amide linkages. The improved asymmetric hairpin provides for the synthesis of tandemly-linked imporoved polyamides that allow for longer binding sites without compromising affinity or selectivity. The improved asymmetric hairpin may also serve to attach functional or detectable groups to the polyamide.

## BRIEF DESCRIPTION OF THE FIGURES

**Figure 1.** A. Hydrogen bonding model of polyamide 1-*R*, ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp, to the DNA sequence 5'-TGTTA-3'. B. Binding model of polyamide 1-*S*, ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp, to the DNA sequence 5'-TGTTA-3'.

**Figure 2.** Computer generated models of: (A) ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp and (B) ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp bound in the minor groove of double stranded DNA van der Waals surface.

**Figure 3.** Structures of the 6-ring hairpin polyamides.

**Figure 4.** Solid phase synthetic scheme for improved polyamides.

**Figure 5.** Results of MPE•Fe(II) footprinting using improved polyamides.

**Figure 6.** Binding patterns of certain improved polyamides to a 135 bp restriction fragment.

**Figure 7.** Affinity cleavage experiments using improved polyamides and a 3'-$^{32}$P-labeled 135 bp restriction fragment.

**Figure 8.** Affinity cleavage patterns of certain improved polyamides at 1 μM concentration and 10 μM concentrations.

**Figure 9.** Panels A-C represent affinity cleavage patterns of certain improved polyamides.

**Figure 10.** Quantitative DNase I footprint titration of certain improved polyamides.

**Figure 11.** Quantitative DNase I footprint titrations of ImPyPy-(R) $^{H_2N}$γ-PyPyPy-β-Dp.

**Figure 12.** Model for chiral hairpin folding of improved polyamides.

**Figure 13.** Hydrogen bonding model of a tandemly-linked polyamide.

**Figure 14.** Structures of exemplary twelve-ring polyamides.

**Figure 15.** Synthesis of tandemly-linked polyamides.

**Figure 16.** Quantitative DNA footprint titrations of an exemplary tandemly-linked polyamide.

**Figure 17.** Exemplary tandemly-linked polyamides.

**Figure 18.** Construction of plasmids pDH10, pDH11, and pDH12.

## DETAILED DESCRIPTION

Within this application, unless otherwise stated, definitions of the terms and illustration of the techniques of this application may be found in any of several well-known references such as: Sambrook, J., *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press (1989); Goeddel, D., *ed.*, *Gene Expression Technology*, *Methods in Enzymology*, **185**, Academic Press, San Diego, CA (1991); "Guide to Protein Purification" in Deutshcer, M.P., ed., *Methods in Enzymology*, Academic Press, San Diego, CA (1989); Innis, et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press, San Diego, CA (1990); Freshney, R.I., *Culture of Animal Cells: A Manual of Basic Technique, 2nd Ed.*, Alan Liss, Inc. New York, NY (1987); Murray, E.J., *ed., Gene Transfer and Expression Protocols*, pp. 109-128, The Humana Press Inc., Clifton, NJ and Lewin, B., *Genes VI,* Oxford University Press, New York (1997).

For the purposes of this application, a *promoter* is a regulatory sequence of DNA that is involved in the binding of RNA polymerase to initiate transcription of a gene. A *gene* is a segment of DNA involved in producing a peptide, polypeptide or protein, including the coding region, non-coding regions preceding ("leader") and following ("trailer") the coding region, as well as intervening non-coding sequences ("introns") between individual coding segments ("exons"). Coding refers to the representation of amino acids, start and stop signals in a three base "triplet" code. Promoters are often upstream (" '5 to") the transcription initiation site of the corresponding gene. Other regulatory sequences of DNA in addition to promoters are known, including sequences involved with the binding of transcription factors, including response elements that are the DNA sequences bound by inducible factors. Enhancers comprise yet another group of regulatory sequences of DNA that can increase the utilization of promoters, and can function in either orientation (5'-3' or 3'-5') and in any location (upstream or downstream) relative to the promoter. Preferably, the regulatory sequence has a positive activity, i.e., binding of an endogeneous ligand (e.g. a transcription factor) to the regulatory sequence increases transcription, thereby resulting in increased expression of the corresponding target gene. In such a case, interference with transcription by binding a polyamide to a regulatory sequence would reduce or abolish expression of a gene.

The promoter may also include or be adjacent to a regulatory sequence known in the art as a *silencer*. A silencer sequence generally has a negative regulatory effect on expression of the gene. In such a case, expression of a gene may be increased directly by using a polyamide to prevent binding of a factor to a silencer regulatory sequence or indirectly, by using a polyamide to block transcription of a factor to a silencer regulatory sequence.

It is to be understood that the polyamides of this invention bind to double stranded DNA in a sequence specific manner. The function of a segment of DNA of a given sequence, such as 5'-TATAAA-3', depends on its position relative to other functional regions in the DNA sequence. In this case, if the sequence 5'-TATAAA-3' on the coding strand of DNA is positioned about 30 base pairs upstream of the transcription start site, the sequence forms part of the promoter region (Lewin, *Genes VI,* pp. 831-835). On the other hand, if the sequence 5'-TATAAA-3' is downstream of the transcription start

site in a coding region and in proper register with the reading frame, the sequence encodes the tyrosyl and lysyl amino acid residues (Lewin, *Genes VI*, pp. 213-215).

While not being held to one hypothesis, it is believed that the binding of the polyamides of this invention modulate gene expression by altering the binding of DNA binding proteins, such as RNA polymerase, transcription factors, TBF, TFIIIB and other proteins. The effect on gene expression of polyamide binding to a segment of double stranded DNA is believed to be related to the function, e.g., promoter, of that segment of DNA.

It is to be understood by one skilled in the art that the improved polyamides of the present invention may bind to any of the above-described DNA sequences or any other sequence having a desired effect upon expression of a gene. In addition, U.S. Patent No. 5,578,444 describes numerous promoter targeting sequences from which base pair sequences for targeting an improved polyamide of the present invention may be identified.

It is generally understood by those skilled in the art that the basic structure of DNA in a living cell includes both *major* and a *minor groove*. For the purposes of describing the present invention, the *minor groove* is the narrow groove of DNA as illustrated in common molecular biology references such as Lewin, B., *Genes VI*, Oxford University Press, New York (1997).

To affect gene expression in a cell, which may include causing an increase or a decrease in gene expression, a effective quantity of one or more polyamide is contacted with the cell and internalized by the cell. The cell may be contacted *in vivo* or *in vitro*. Effective extracellular concentrations of polyamides that can modulate gene expression range from about 10 nanomolar to about 1 micromolar. Gottesfeld, J.M., *et al.*, *Nature* **387** 202-205 (1997). To determine effective amounts and concentrations of polyamides *in vitro*, a suitable number of cells is plated on tissue culture plates and various quantities of one or more polyamide are added to separate wells. Gene expression following exposure to a polyamide can be monitored in the cells or medium by detecting the amount of the protein gene product present as determined by various techniques utilizing specific antibodies, including ELISA and western blot. Alternatively, gene expression following

exposure to a polyamide can be monitored by detecting the amount of messenger RNA present as determined by various techniques, including northern blot and RT-PCR.

Similarly, to determine effective amounts and concentrations of polyamides for *in vivo* administration, a sample of body tissue or fluid, such as plasma, blood, urine, cerebrospinal fluid, saliva, or biopsy of skin, muscle, liver, brain or other appropriate tissue source is analyzed. Gene expression following exposure to a polyamide can be monitored by detecting the amount of the protein gene product present as determined by various techniques utilizing specific antibodies, including ELISA and western blot. Alternatively, gene expression following exposure to a polyamide can be monitored by the detecting the amount of messenger RNA present as determined by various techniques, including northern blot and RT-PCR.

The polyamides of this invention may be formulated into diagnostic and therapeutic compositions for *in vivo* or *in vitro* use. Representative methods of formulation may be found in *Remington: The Science and Practice of Pharmacy*, 19th ed., Mack Publishing Co., Easton, PA (1995).

For *in vivo* use, the polyamides may be incorporated into a physiologically acceptable pharmaceutical composition that is administered to a patient in need of treatment or an animal for medical or research purposes. The polyamide composition comprises pharmaceutically acceptable carriers, excipients, adjuvants, stabilizers, and vehicles. The composition may be in solid, liquid, gel, or aerosol form. The polyamide composition of the present invention may be administered in various dosage forms orally, parentally, by inhalation spray, rectally, or topically. The term parenteral as used herein includes, subcutaneous, intravenous, intramuscular, intrasternal, infusion techniques or intraperitoneally.

The selection of the precise concentration, composition, and delivery regimen is influenced by, *inter alia*, the specific pharmacological properties of the particular selected compound, the intended use, the nature and severity of the condition being treated or diagnosed, the age, weight, gender, physical condition and mental acuity of the intended recipient as well as the route of administration. Such considerations are within the purview of the skilled artisan. Thus, the dosage regimen may vary widely, but can be determined routinely using standard methods.

Polyamides of the present invention are also useful for detecting the presence of double stranded DNA of a specific sequence for diagnostic or preparative purposes. The sample containing the double stranded DNA can be contacted by polyamide linked to a solid substrate, thereby isolating DNA comprising a desired sequence. Alternatively, polyamides linked to a suitable detectable marker, such as biotin, a hapten, a radioisotope or a dye molecule, can be contacted by a sample containing double stranded DNA.

The design of bifunctional sequence specific DNA binding molecules requires the integration of two separate entities: recognition and functional activity. Polyamides that specifically bind with subnanomolar affinity to the minor groove of a predetermined sequence of double stranded DNA are linked to a functional molecule, providing the corresponding bifunctional conjugates useful in molecular biology, genomic sequencing, and human medicine. Polyamides of this invention can be conjugated to a variety of functional molecules, which can be independently chosen from but is not limited to arylboronic acids, biotins, polyhistidines comprised from about 2 to 8 amino acids, haptens to which an antibody binds, solid phase supports, oligodeoxynucleotides, N-ethylnitrosourea, fluorescein, bromoacetamide, iodoacetamide, DL-α-lipoic acid, acridine, captothesin, pyrene, mitomycin, texas red, anthracene, anthrinilic acid, avidin, DAPI, isosulfan blue, malachite green, psoralen, ethyl red, 4-(psoraen-8-yloxy)-butyrate, tartaric acid, (+)-α-tocopheral, psoralen, EDTA, methidium, acridine, Ni(II)•Gly-Gly-His, thiazole orange (TO), Dansyl, pyrene, N-bromoacetamide, and gold particles. Such bifunctional polyamides are useful for DNA affinity capture, covalent DNA modification, oxidative DNA cleavage, DNA photocleavage. Such bifunctional polyamides are useful for DNA detection by providing a polyamide linked to a detectable label. Detailed instructions for synthesis of such bifunctional polyamides can be found in copending U.S. provisional application 60/043,444, the teachings of which are incorporated by reference.

DNA complexed to a labeled polyamide can then be determined using the appropriate detection system as is well known to one skilled in the art. For example, DNA associated with a polyamide linked to biotin can be detected by a streptavidin / alkaline phosphatase system.

The present invention also describes a diagnostic system, preferably in kit form, for assaying for the presence of the double stranded DNA sequence bound by the

polyamide of this invention in a body sample, such brain tissue, cell suspensions or tissue sections, or body fluid samples such as CSF, blood, plasma or serum, where it is desirable to detect the presence, and preferably the amount, of the double stranded DNA sequence bound by the polyamide in the sample according to the diagnostic methods described herein.

The diagnostic system includes, in an amount sufficient to perform at least one assay, a specific polyamide as a separately packaged reagent. Instructions for use of the packaged reagent(s) are also typically included. As used herein, the term "package" refers to a solid matrix or material such as glass, plastic (e.g., polyethylene, polypropylene or polycarbonate), paper, foil and the like capable of holding within fixed limits a polyamide of the present invention. Thus, for example, a package can be a glass vial used to contain milligram quantities of a contemplated polyamide or it can be a microliter plate well to which microgram quantities of a contemplated polypamide have been operatively affixed, i.e., linked so as to be capable of being bound by the target DNA sequence. "Instructions for use" typically include a tangible expression describing the reagent concentration or at least one assay method parameter such as the relative amounts of reagent and sample to be admixed, maintenance time periods for reagent or sample admixtures, temperature, buffer conditions and the like. A diagnostic system of the present invention preferably also includes a detectable label and a detecting or indicating means capable of signaling the binding of the contemplated polyamide of the present invention to the target DNA sequence. As noted above, numerous detectable labels, such as biotin, and detecting or indicating means, such as enzyme-linked (direct or indirect) streptavidin, are well known in the art.

Trauger,et al. (*Nature*, 382: 559-561) and Swalley, et al. (*J. Am. Chem. Soc.* 119: 6953-6961) have described recognition of DNA by certain polyamides at subnanomolar concentrations. Pairing specific carboxyamide groups allows for recognition of specific DNA sequences (Swalley, et al. *supra*). Polyamides comprising Hp, Im, and Py provide for coded targeting of pre-determined DNA sequences with high affinity and specificity. Hp, Im, and Py polyamides may be combined to form Im/Py, Py/Im, Hp/Py, and Py/Hp binding pairs which complement the four Watson-Crick base pairs A, C, G, and T. Table 1 illustrates such pairings.

**TABLE 1**

*Pairing Codes for Base Pair Recognition**

| Pair | G•C | C•G | T•A | A•T |
|---|---|---|---|---|
| Im/Py | + | - | - | - |
| Py/Im | - | + | - | - |
| Hp/Py | - | - | + | - |
| Py/Hp | - | - | - | + |

*favored (+), disfavored (-)

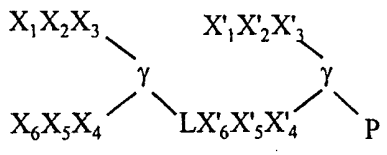
Three-, four-, five- or six-ring improved polyamides of the present invention are covalently coupled to form six-, eight-, ten- or twelve-ring strucures, respectively, that bind specifically to four or six base pair targets, respectively, at subnanomolar concentrations. As such, the improved polyamides of the present invention may be directed to any DNA sequence comprised of A, C, G, or T.

The improved polyamides of the present invention comprise those having at least three consecutive carboxamide pairings for binding DNA in the minor groove of a regulatory sequence of a duplex gene sequence and a chiral hairpin turn with a stereochemical center substituted at the γ-position of the chiral hairpin turn of the molecule with the R-enantiomer of 2,4-diaminobutyric acid ($H_2NHCHCH_2CHNH_2$-COOH; "$(R)^{H_2N}\gamma$"). In addition, the present invention provides a methodology for covalently linking existing hairpin motifs without compromising DNA-binding and sequence specificity. The present invention provides improved polyamides for binding the minor groove of DNA to affect gene expression. Preferably, the bound polyamide inhibits gene expression.

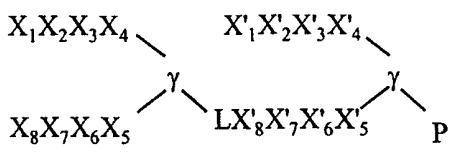
The present invention comprises improved polyamides having three or four-ring polyamide structures covalently coupled to form six-, eight-, ten- or twelve-ring hairpin structures, respectively, of the general structures I-VIII:

$X_1X_2X_3\gamma X_4X_5X_6$ **I** $X_1X_2X_3X_4\gamma X_5X_6X_7X_8$ **II** $X_1X_2X_3X_4X_5\gamma X_6X_7X_8X_9X_{10}$ **III**

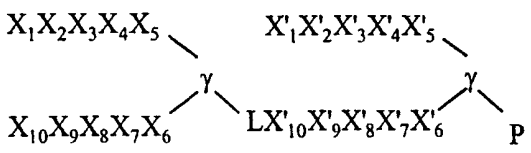
$X_1X_2X_3X_4X_5X_6\gamma X_7X_8X_9X_{10}X_{11}X_{12}$ **IV**



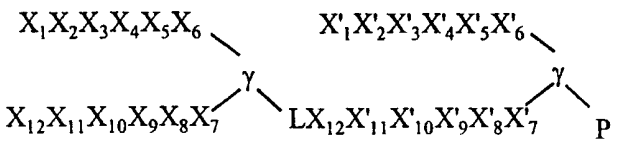

V




VI




VII




VIII

where $X_{1-12}$ and $X'_{1-12}$ are independently an imidazole such as N-methylimidazole (Im), a pyrrole such as N-methylpyrrole (Py), or a hydroxypyrrole such as 3-hydroxy-N-methyl pyrrole (Hp). In addition, an improved polyamide of the present invention may further include a aliphatic amino acid such as β-alanine residue (β), an amide group such as dimethylaminopropylamide (Dp), an alcohol such as EtOH, an acid such as EDTA, or any derivative thereof may be joined to the β residue.

β-alanine may also be utilized in place of a pyrrole amino acid in Formulas I-VIII. The use of β-alanine in place of a pyrrole or hydroxypyrrole amino acid in the synthetic methods provides aromatic/aliphatic pairing (Im/β, β/Im, Py/β, and β/Py) and aliphatic/aliphatic pairing (β/β) substitution. Such substitutions may comprise those described in provisional application 60/042,022, incorporated herein by reference. The use of γ-aminobutyric acid, or a substituted γ-aminobutyric acid such as (R)-2,4 diaminobutyric acid, provides for preferred hairpin turns. Many other groups suitable for the purposes of practicing this invention are well known and widely available to one skilled in the art.

The polyamide subunit structures I-VIII above are covalently coupled through the γ residue which represents a -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid. The present invention provides the reagents and methodologies for substituting the γ-residue of certain polyamides with a moiety such as (*R*)-2,4,-diaminobutyric acid ($(R)^{H_2N}\gamma$). The NMR structure of a hairpin polyamide of sequence composition ImPyPy-γ-PyPyPy complexed with a 5'-TGTTA-3' target site indicated that it was possible to substitute the α-position of the γ-aminobutyric acid residue within the hairpin-DNA complex (de Claire, et al. *J. Am. Chem. Soc.* 1997, *119*, 7909). Modeling indicated that replacing the α-H of γ with an amino group that may confer an *R*-configuration at the α-carbon could be accommodated within the floor and walls of the minor groove as demonstrated in Figure 1 and 2A. In contrast, the (*S*)-2,4,-diaminobutyric acid ($(S)^{H_2N}\gamma$) linked hairpin is predicted to clash with the walls of the minor groove of the DNA helix as illustrated in Figures 1 and 2B.

In Formulas V-VIII, L represents an amino acid linking group such as β-alanine or 5-aminovaleric acid (δ) bound to the γ residue of a first polyamide and to the carboxytail of a second polyamide. As such, two or more polyamides may be linked, forming a tandemly-linked polyamide. Such a polyamide is said to be tandemly-linked or a tandem-linked polyamide.

P represents from zero to ten polyamides of formulas I-VIII that may be tandemly linked to the second polyamide. Preferably, P represents from zero to eight polyamids of formulas I-VIII. More preferably, P represents from zero to six polyamids of formulas I-VIII. More preferably, P represents from zero to four polyamides of formulas I-VIII. Most preferably, P represents from zero to two polyamides of formulas I-VIII. Tandem linking of polyamides provides expanded binding site size and increased binding affinity without compromising selectivity. Many other groups suitable for the purposes of practicing this invention are well known and widely available to one skilled in the art.

Baird, et al. (*J. Am. Chem. Soc.* 118: 6141-6146) and PCT/US97/003332 describe methods for synthesis of polyamides which are suitable for preparing polyamides of this invention. Polyamides of the present invention may be synthesized by solid phase

methods using compounds such as Boc-protected 3-methoxypyrrole, imidazole, and pyrrole aromatic amino acids, which are cleaved from the support by aminolysis, deprotected with sodium thiophenoxide, and purified by reverse-phase HPLC. The identity and purity of the polyamides may be verified using any of a variety of analytical techniques available to one skilled in the art such as 1H-NMR, analytical HPLC, and/or matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS-monoisotropic).

Described herein is the synthesis of a new class of chiral hairpin polyamides and their characterization with regard to DNA binding affinity and sequence specificity. The present invention provides one skilled in the art with the reagents and methodologies for substitution of the prochiral γ-turn with either enantiomer of 2,4-diaminobutyric acid. In addition, the invention provides the dicationic six-ring enantiomeric polyamides (+)-ImPyPy-$(R)^{H_2N}$ γ-PyPyPy-β-Dp (1-*R*) and (-)-ImPyPy-$(S)^{H_2N}$ γ-PyPyPy-β-Dp (1-*S*) which may be synthesized by solid phase methods. In certain experiments, the monocationic polyamide (+)-ImPyPy-$(R)^{H_2N}$ γ PyPyPy-β-EtOH (2-*R*), which lacks a charge at the C-terminus, may be prepared and utilized as a control. To further study steric effects, the γ-acetamido polyamides (+)-ImPyPy-$(R)^{Ac}$ γ -PyPyPy-β-Dp (3-*R*) and (-)-ImPyPy-$(S)^{Ac}$ γ-PyPyPy-β-Dp (3-*S*) may be utilized (Figure 3; Baird, et al. 1996. *J. Am. Chem. Soc.* 118: 6141). The present invention further provides the EDTA analogs ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp-EDTA•Fe(II) (4-*R*•Fe(II)), ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp-EDTA•Fe(II) (4-*S*•Fe(II)), ImPyPy-$(R)^{EDTA\cdot Fe(II)}$ γ-PyPyPy-β-Dp (5-*R*•Fe(II)), and ImPyPy-$(S)^{EDTA\cdot Fe(II)}$ γ-PyPyPy-β-Dp (5-*S*•Fe(II)) that may be utilized to to confirm the binding orientation of the modified hairpins at specific DNA binding sites (Figure 3).

Tandemly-linked polyamides of the present invention are also provided by the instant invention. The primary turn-amino group provides a potential site for covalently tethering two hairpins. In one potential linkage arrangement, the C-terminus of the first hairpin is coupled to the α-amino group of the γ-turn of the second amino acid linker. The present invention provides twelve ting polyamides exemplified by ImPyPy-(*R*)[ImPyPy-$(R)^{H_2N}$γPyPyPy-β]$^{HN}$γPyPyPy-β-Dp and ImPyPy-(*R*)[ImPyPy-$(R)^{H_2N}$γPyPyPy-δ]$^{HN}$γPyPyPy-β-Dp (Figure 14). The DNA binding properties of certain

polyamides of the present invention were determined on a series of DNA fragments containing 10, 11 and 12 base pair target sites. The present invention further provides an exemplary affinity cleaving derivative ImPyPy-(*R*)[ImPyPy-(*R*)$^{EDTA}$γPyPyPy-δ]$^{HN}$γPyPyPy-β-Dp, which was utilized to confirm a single predicted binding orientation for the tandemly-linked polyamide. Methodologies for the determination of the DNA-binding affinity and sequence selectivity of tandem improved polyamides is also provided.

The present invention reveals to one skilled in the art properties of chiral structure elements that may be utilized as a guide in the design of more efficient polyamides. For instance, the present invention provides amine substituents on the $(R)^{H_2N}\gamma$ turn amino acid that enhance the DNA binding affinity and specificity relative to the unsubstituted parent hairpin, providing for an optimized class of hairpin polyamides. Also provided are acetamido substituents at the $(R)^{H_2N}\gamma$ that do not compromise affinity or specificity relative to the parent hairpin, providing for a convenient synthetic attachment point at the 'capped' end of the molecule. In addition, the invention described herein provides $(S)^{H_2N}\gamma$-linked hairpins that bind with enhanced affinity to reverse orientation sites relative to the parent hairpin and $(R)^{H_2N}\gamma$-linked hairpins that bind with enhanced specificity relative to the parent hairpin indicating that γ-turn substituents may regulate hairpin polyamide binding orientational preference. The invention further provides the skilled artisan with the necessary tools and methodologies for developing tandemly-linked polyamides to increase the polyamide binding site size, and increase affinity without compromising sequence selectivity.

The examples listed above and those illustrated below represent only certain embodiments of the present invention and are not limiting of the specification and claims in any way.

## EXAMPLES

### Example 1

***Synthesis of Improved Polyamides***

Two polyamide-resins ImPyPy-$(R)^{Fmoc}$ γ-PyPyPy-β-Pam-resin and ImPyPy-$(S)^{Fmoc}$ γ-PyPyPy-β-Pam-resin, were synthesized in 14 steps from Boc-β-alanine-Pam-resin (1 g resin, 0.2 mmol/g substitution) using previously described Boc-chemistry machine-assisted protocols (Figure 4; Baird, et al. *J. Am. Chem. Soc.* **1996**, *118*, 6141). (*R*)- and (*S*)-2,4-diaminobutyric acid residues were introduced as orthogonally protected *N*- γ-Fmoc-*N*- γ-Boc derivatives (HBTU, DIEA). Fmoc protected polyamide resins ImPyPy-$(R)^{Fmoc}$ γ-PyPyPy-β-Pam-resin and ImPyPy-$(S)^{Fmoc}$ γ-PyPyPy-β-Pam-resin were treated with 1:4 DMF:Piperidine (22 °C, 30 min.) to provide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Pam-resin and ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Pam-resin, respectively. A single-step aminolysis of the resin ester linkage was used to cleave the polyamide from the solid support. A sample of resin (240 mg) was treated with either dimethylaminopropylamine (55 °C, 18 h) to provide **1-*R***, **1-*S***, **3-*R***, and **3-*S*** or ethanolamine (55 °C, 18 h) to provide **2-*R***. Resin cleavage products were purified by reverse phase HPLC to provide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp (**1-*R***), ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp (**1-*S***), and ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-EtOH (**2-*R***). The stereochemical purity of **1-*R*** was determined to be > 98% by Mosher amide analysis (Dale, et al. *J. Am. Chem. Soc.* **1973**, *95*, 512; Yamaguchi, et al. *Asymmetric Synthesis (Vol. 1), Analytical Methods* p.125-152, J. D. Morrison (ed.) Academic Press (1983)). **1-*R,R*** and **1-*R,S*** Mosher amides were prepared by reaction of **1-*R*** with HOBt activated esters generated *in situ* from (*R*)- α-methoxy-α-(triflouromethyl)phenylacetic acid and (*S*)- α-methoxy-α-(triflouromethyl)phenylacetic acid. For synthesis of analogs modified with EDTA at the carboxy-terminus, the amine-resin was treated with Boc-anhydride (DMF, DIEA, 55 °C, 30 min) to provide ImPyPy-$(R)^{Boc}$γPyPyPy-β-Pam-resin and ImPyPy-$(S)^{Boc}$γ–PyPyPy-β-Pam-resin (Figure 4). A sample of Boc-resin was then cleaved with 3,3'-diamino-*N*-methyldipropylamine (55 °C,

18 h) and purified by reversed phase HPLC to provide either ImPyPy-$(R)^{Boc}$γ-PyPyPy-β-Dp-$NH_2$ (**1-*R*-Boc-$NH_2$**) or ImPyPy-$(S)^{Boc}$γ–PyPyPy-β-Dp-$NH_2$ (**1-*S*-Boc-$NH_2$**) which afford free primary amine groups at the C-terminus suitable for post-synthetic modification. The polyamide-amines **1-*R*-Boc-$NH_2$** and **1-*S*-Boc-$NH_2$** were treated with an excess of EDTA-dianhydride (DMSO/NMP, DIEA, 55 °C, 15 min) and the remaining anhydride hydrolyzed (0.1 M NaOH, 55 °C, 10 min). The Boc protected EDTA modified polyamides ImPyPy-$(R)^{Boc}$γ–PyPyPy-β-Dp-EDTA (**4-*R*-Boc**) and ImPyPy-$(S)^{Boc}$γ-PyPyPy-β-Dp-EDTA (**4-*S*-Boc**) were isolated by HPLC. Individual Boc-EDTA-polyamides were deprotected with neat TFA (22 °C, 1 h) to provide the respective C-terminal EDTA derivatives, ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp-EDTA (**4-*R***) and ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp-EDTA (**4-*S***). For the synthesis of acetamide-turn or EDTA-turn derivatives, a sample of the γ-amino polyamide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp (**1-*R***) or ImPyPy-$(S)^{H_2N}$γ–PyPyPy-β-Dp (**1-*S***) was treated with an excess of either acetic anhydride or EDTA-dianhydride (DMSO/NMP, DIEA 55 °C, 30 min) and the remaining anhydride hydrolyzed (0.1 M NaOH, 55 °C, 10 min). The polyamides ImPyPy-$(R)^{Ac}$γ–PyPyPy-β-Dp (**3-*R***), ImPyPy-$(S)^{Ac}$γ–PyPyPy-β–Dp (**3-*S***), ImPyPy-$(R)^{EDTA}$γ-PyPyPy-β-Dp (**5-*R***) and ImPyPy-$(S)^{EDTA}$γ-PyPyPy-β-Dp (**5-*S***) were then isolated by reverse phase HPLC. The six-ring hairpin polyamides described here are soluble in aqueous solution at concentrations 10 mM at 37°C.

**A. Materials**

Dicyclohexylcarbodiimide (DCC), Hydroxybenzotriazole (HOBt), 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexa-fluorophosphate (HBTU) and 0.2 mmol/gram Boc-β-alanine-(-4-carboxamidomethyl)-benzyl-ester-copoly(styrene-divinylbenzene) resin (Boc-γ-Pam-Resin) was purchased from Peptides International (0.2 mmol/gram) (*R*)-2-Fmoc-4-Boc-diaminobutyric acid, (*S*)-2-Fmoc-4-Boc-diaminobutyric acid, and (*R*)-2-amino-4-Boc-diaminobutyric acid were from Bachem. *N,N*-diisopropylethylamine (DIEA), *N,N*-dimethylformamide (DMF), *N*-methylpyrrolidone

(NMP), DMSO/NMP, Acetic anhydride ($Ac_2O$), and 0.0002 M potassium cyanide/pyridine were purchased from Applied Biosystems. Dichloromethane (DCM) and triethylamine (TEA) were reagent grade from EM, thiophenol (PhSH), dimethylaminopropylamine (Dp), (*R*)- α-methoxy-α-(trifuoromethyl)phenylacetic acid ((*R*)MPTA) and (*S*)-α-methoxy-α-(trifouromethyl)phenylacetic acid ((*S*)MPTA) were from Aldrich, trifluoroacetic acid (TFA) Biograde from Halocarbon, phenol from Fisher, and ninhydrin from Pierce. All reagents were used without further purification.

Quik-Sep polypropylene disposable filters were purchased from Isolab Inc. A shaker for manual solid phase synthesis was obtained from St. John Associates, Inc. Screw-cap glass peptide synthesis reaction vessels (5 mL and 20 mL) with a #2 sintered glass frit were made as described by Kent (*Annu. Rev. Biochem.* **1988**, *57*, 957). $^1$H NMR spectra were recorded on a General Electric-QE NMR spectrometer at 300 MHz with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured in water on a Hewlett-Packard Model 8452A diode array spectrophotometer. Optical rotations were recorded on a JASCO Dip 1000 Digital Polarimeter. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF) was performed at the Protein and Peptide Microanalytical Facility at the California Institute of Technology. HPLC analysis was performed on either a HP 1090M analytical HPLC or a Beckman Gold system using a RAINEN $C_{18}$, Microsorb MV, 5μm, 300 x 4.6 mm reversed phase column in 0.1% (wt/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory reverse phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C18 column equipped with a guard, 0.1% (wt/v) TFA, 0.25% acetonitrile/min. Distilled water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered.

Enzymes were purchased from Boehringer-Mannheim and used with their supplied buffers. Deoxyadenosine and thymidine 5′-[$\gamma^{32}$P] triphosphates were obtained

from Amersham, and deoxyadenosine 5′-[$\gamma^{32}$P]triphosphate was purchased from I.C.N. Sonicated, deproteinized calf thymus DNA was acquired from Pharmacia. RNase free water was obtained from USB and used for all footprinting reactions. All other reagents and materials were used as received. All DNA manipulations were performed according to standard protocols (Sambrook, J.; Fritsch, E.F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989).

**B. ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Dp (1-*R*)**

ImPyPy-(*R*)$^{Fmoc}$γ-PyPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods (Baird, et al. *J. Am. Chem. Soc.* **1996**, *118*, 6141). (*R*)-2-Fmoc-4-Boc-diaminobutyric acid (0.7 mmol) was incorporated as previously described for Boc-γ-aminobutyric acid. ImPyPy-(*R*)$^{Fmoc}$γ–PyPyPy-β-Pam-Resin was placed in a glass 20 mL peptide synthesis vessel and treated with DMF (2 mL), followed by piperidine (8 mL) and agitated (22 °C, 30 min.). ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Pam-resin was isolated by filtration, and washed sequentially with an excess of DMF, DCM, MeOH, and ethyl ether and the amine-resin dried *in vacuo*. A sample of ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Pam-resin (240 mg, 0.18 mmol/gram) was treated with neat dimethylaminopropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. Resin substitution can be calculated as $L_{new}(mmol/g) = L_{old}/(1 + L_{old}(W_{new} - W_{old}) \times 10^{-3})$, where L is the loading (mmol of amine per gram of resin), and W is the weight ($gmol^{-1}$) of the growing polyamide attached to the resin (Barlos, et al. *Int. J. Peptide Protein Res.* **1991**, *37*, 513). The reaction mixture was placed in an oven and periodically agitated (55 °C, 16 h). The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (32 mg, 66% recovery). $[\alpha]^{20}_D$ +14.6 (*c* 0.05, $H_2O$); UV ($H_2O$)$\lambda_{max}$ 246, 310 (50,000); $^1$H NMR (DMSO-$d_6$) 10.56 (s, 1 H), 10.47 (s, 1 H), 9.97 (s,

1 H), 9.94 (s, 1 H), 9.88 (s, 1 H), 9.4 (br s, 1 H), 8.28 (s, 3 H), 8.22 (m, 1 H), 8.03 (m, 2 H), 7.38 (s, 1 H), 7.25 (d, 1 H, $J$ = 1.6 Hz), 7.22 (d, 1 H, $J$ = 1.5 Hz), 7.19 (d, 1 H, $J$ = 1.5 Hz), 7.16 (d, 1 H, $J$ = 1.6 Hz), 7.14 (d, 1 H, $J$ = 1.8 Hz), 7.12 (d, 1 H, $J$ = 1.7 Hz), 7.03 (m, 2 H), 6.95 (d, 1 H, J = 1.6 Hz), 6.91 (d, 1 H, $J$ = 1.6 Hz), 6.85 (d, 1 H, $J$ = 1.6 Hz), 3.96 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.76 (s, 3 H), 3.33 (q, 2 H, $J$ = 6.3 Hz), 3.25 (q, 2 H, $J$ = 5.7 Hz), 3.05 (q, 2 H, $J$ = 5.9 Hz), 2.96 (q, 2 H, $J$ = 5.3 Hz), 2.71 (d, 6 H, $J$ = 4.9 Hz), 2.32 (t, 2 H, $J$ = 7.1 Hz), 1.95 (q, 2 H, $J$ = 5.9 Hz), 1.70 (quintet, 2 H, $J$ = 7.3 Hz); MALDI-TOF-MS (monoisotopic), 992.5 (992.5 calc. for $C_{47}H_{62}N_{17}O_8$).

A hydrogen bonding model of the 1:1 polyamide:DNA complex formed between the hairpin polyamide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp (**1-*R***) with a 5'-TGTTA-3' site is illustrated in Figure 1A. Lone pairs of $N_3$ of purines and $O_2$ of pyrimidines are shown as circles with dots. The $N_2$ hydrogen of guanine is indicated by circles containing an "H" and putative hydrogen bonds are illustrated by dotted lines. A schematic binding model is also demonstrated where the imidazole and pyrrole rings are represented as shaded and unshaded spheres, respectively, and the β-alanine residue is represented as an unshaded diamond. Figure 2 shows models derived from the NMR structure coordinates of ImPyPy-γ-PyPyPy-β-Dp•5'-TGTTA-3'using InsightII software. Figure 3 shows the structure of the *2-R* polyamides.

### C. ImPyPy-$(R)^{(R)MTPA}$γ-PyPyPy-β-Dp (1-*R,R*)

(*R*)-α-methoxy-α-(triflouromethyl)phenylacetic acid (117 mg, 0.5 mmol) and HOBt (70 mg, 0.5 mmol) were dissolved in DMF (1 mL), DCC (100 mg, 0.5 mmol) added and the solution agitated for 30 min at 22 °C. A sample of the activated ester solution (100 μL, 0.05 mmol) was added to ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp **1-*R*** (10 mg, 0.01 mmol), DIEA (50 μL) added, and the solution agitated for 3h (22 °C). DMF (1 mL) followed by 0.1% (wt/v) TFA (6 mL) was then added to the reaction mixture and the resulting solution purified by reversed phase HPLC (1% acetonitrile/min.) under

conditions which were determined to separate the diastereomers. ImPyPy-$(R)^{(R)MTPA}$γ-PyPyPy-β-Dp is recovered as a white powder upon lyophilization of the appropriate fractions (6 mg, 53% recovery). $^1$H NMR (DMSO-$d_6$) δ10.50 (s, 1 H), 10.14 (s, 1 H), 9.92 (s, 2 H), 9.88 (s, 1 H), 9.2 (br s, 1 H), 8.43 (d, 1 H, $J$ = 7.0 Hz), 8.02 (m, 3 H), 7.92 (m, 1 H), 7.47 (m, 2 H), 7.41 (m, 2 H), 7.36 (s, 1 H), 7.24 (m, 1 H), 7.19 (m, 1 H), 7.15 (m, 1 H), 7.12 (m, 3 H), 7.01 (m, 2 H), 6.90 (m, 3 H), 6.83 (m, 1 H), 4.46 (q, 1 H, $J$ = 5.5 Hz), 3.94 (s, 3 H), 3.79 (m, 9 H), 3.75 (m, 6 H), 3.32 (m, 4 H), 3.05 (m, 2 H), 2.94 (m, 2 H), 2.68 (d, 6H, $J$ = 4.0 Hz), 2.28 (t, 2 H, $J$ = 6.3 Hz), 1.93 (q, 2 H, $J$ = 6.1 Hz), 1.66 (quintet, 2 H, $J$ = 6.0 Hz), 1.18 (s, 3 H); MALDI-TOF-MS (monoisotopic), 1208.5 (1208.5 calc. for $C_{57}H_{68}F_3N_{17}O_{10}$).

**D. ImPyPy-$(R)^{(S)MTPA}$γ-PyPyPy-β-Dp 1-*R,S***

ImPyPy-$(R)^{(S)MTPA}$γ-PyPyPy-β-Dp was prepared from (*S*)-α-methoxy-α-(triflouromethyl)phenylacetic acid as described for **1-*R,R*** (5 mg, 45% recovery). $^1$H NMR (DMSO-$d_6$) δ10.47 (s, 1 H), 10.08 (s, 1 H), 9.92 (s, 2 H), 9.88 (s, 1 H), 9.2 (br s, 1 H), 8.43 (d, 1 H, $J$ = 6.9 Hz), 8.02 (m, 3 H), 7.46 (m, 2 H), 7.40 (m, 2 H), 7.36 (s, 1 H), 7.23 (m, 1 H), 7.19 (m, 1 H), 7.14 (m, 1 H), 7.12 (m, 3 H), 7.01 (m, 2 H), 6.87 (m, 3 H), 6.83 (m, 1 H), 4.44 (q, 1 H, $J$ = 6.5 Hz), 3.94 (s, 3 H), 3.79 (m, 9 H), 3.75 (m, 6 H), 3.28 (m, 4 H), 3.06 (m, 4 H), 2.94 (m, 2 H), 2.69 (d, 6H, $J$ = 4.5 Hz), 2.28 (t, 2 H, $J$ = 6.5 Hz), 1.93 (q, 2 H, $J$ = 6.1 Hz), 1.66 (quintet, 2 H, $J$ = 6.0 Hz), 1.18 (s, 3 H); MALDI-TOF-MS (monoisotopic), 1209.0 (1208.5 calc. for $C_{57}H_{68}F_3N_{17}O_{10}$).

**E. ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp (1-*S*)**

ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp was prepared as described for **1-*R*** (23 mg, 49% recovery). $[\alpha]^{20}_D$ -14.2 (*c* 0.04, $H_2O$); $^1$H NMR (DMSO-$d_6$) identical to **1-*R***; MALDI-TOF-MS (monoisotopic), 992.5 (992.5 calc. for $C_{47}H_{62}N_{17}O_8$). Figure 2B illustrates a binding model for the ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp (1-*S*) polyamide to the DNA seuqnce 5'-TGTTA-3'.

**F. ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-EtOH (2-*R*)**

A sample of ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Pam-resin (240 mg, 0.18 mmol/gram) was treated with neat ethanolamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC to provide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-EtOH as a white powder upon lyophilization of the appropriate fractions (21 mg, 46% recovery). $[\alpha]^{20}_D$ +18.6 (*c* 0.04, $H_2O$); UV ($H_2O$) $\lambda_{max}$ 246, 310 (50,000); $^1$H NMR (DMSO-$d_6$) δ10.55 (s, 1 H), 10.48 (s, 1 H), 9.97 (s, 1 H), 9.94 (s, 1 H), 9.89 (s, 1 H), 8.24 (m, 4 H), 8.00 (t, 1 H, *J* = 4.1 Hz), 7.89 (t, 1 H, *J* = 5.8 Hz), 7.38 (s, 1 H), 7.25 (d, 1 H, *J* = 1.6 Hz), 7.22 (d, 1 H, *J* = 1.6 Hz), 7.21 (d, 1 H, *J* = 1.5 Hz), 7.16 (m, 2 H), 7.14 (d, 1 H, *J* = 1.6 Hz), 7.03 (d, 1 H, *J* = 1.7 Hz), 6.99 (d, 1 H, *J* = 1.4 Hz), 6.95 (d, 1 H, *J* = 1.6 Hz), 6.91 (d, 1 H, J = 1.5 Hz), 6.78 (d, 1 H, *J* = 1.5 Hz), 5.33 (m, 1 H), 3.95 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.76 (s, 3 H), 3.37 (q, 2 H, *J* = 6.2 Hz), 3.07 (q, 2 H, *J* = 5.9 Hz), 2.29 (t, 2 H, *J* = 7.1 Hz), 1.93 (q, 2 H, *J* = 5.8 Hz), 1.20 (m, 4 H); MALDI-TOF-MS (monoisotopic), 951.4 (951.4 calc. for $C_{44}H_{55}N_{16}O_9$). The *2-R* polyamide is shown in Figure 3.

### G. ImPyPy-$(R)^{Ac}$γ-PyPyPy-β-Dp (3-*R*)

A sample of ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp (4 mg) in DMSO (1 mL) was treated with a solution of acetic anhydride (1 mL) and DIEA (1 mL) in DMF (1 ml) and heated (55 °C) with periodic agitation for 30 min. Residual acetic anhydride was hydrolyzed (0.1 M NaOH, 1 mL, 55 °C, 10 min.), 0.1% (wt/v) TFA was added (6 mL) and the resulting solution purified by reversed phase HPLC to provide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp is recovered as a white powder upon lyophilization of the appropriate fractions (2 mg, 50% recovery). $[\alpha]^{20}_D$ +20.5 (*c* 0.06, $H_2O$); UV ($H_2O$) $\lambda_{max}$ 242, 304 (50,000); $^1$H NMR (DMSO-$d_6$) δ10.49 (s, 1 H), 10.06 (s, 1 H), 9.94 (m, 2 H), 9.00 (s, 1 H), 9.4 (br s, 1 H), 8.21 (d, 1 H, *J* = 7.8 Hz), 8.06 (m, 2 H), 8.00 (t, 1 H, *J* = 6.2 Hz), 7.39 (s, 1 H), 7.27 (d, 1 H, *J* = 1.7 Hz), 7.21 (d, 1 H, *J* = 1.6 Hz), 7.18 (m, 2 H), 7.14 (m, 2 H), 7.03 (m, 2 H), 6.90 (d, 1 H, *J* = 1.6 Hz), 6.86 (m, 2 H), 4.43 (q, 1 H, *J* = 7.5 Hz), 3.96 (s, 3 H), 3.82

(m, 9 H), 3.73 (m, 6 H), 3.37 (q, 2 H, $J$ = 5.8 Hz), 3.11 (q, 2 H, $J$ = 6.9 Hz), 2.98 (q, 2 H, $J$ = 5.4 Hz), 2.79 (q, 2 H, $J$ = 5.3 Hz), 2.71 (d, 6 H, $J$ = 4.7 Hz), 2.33 (t, 2 H, $J$ = 6.2 Hz), 1.97 (s, 3 H), 1.70 (quintet, 2 H, $J$ = 6.0 Hz) MALDI-TOF-MS (average), 1035.1 (1035.2 calc. for M+H). The *3-R* polyamide is shown in Figure 3.

5

**H. ImPyPy-(*S*)$^{Ac}$γPyPyPy-β-Dp (3-*S*)**

ImPyPy-(*S*)$^{Ac}$γPyPyPy-β-Dp was prepared as described for **3-*R***. (2 mg, 50% recovery). $[\alpha]^{20}_D$ -16.4 (*c* 0.07, $H_2O$); $^1$H NMR (DMSO-$d_6$) is identical to **3-*R***; MALDI-TOF-MS (monoisotopic), 1034.6 (1034.5 calc. for $C_{49}H_{64}N_{17}O_9$).

**I. ImPyPy-(*R*)$^{Boc}$γPyPyPy-β-Dp-NH$_2$ (4-*R*-Boc-NH$_2$)**

A sample of ImPyPy-(*R*)$^{H_2N}$γPyPyPy-β-Pam-resin (300 mg, 0.18 mmol/gram). Resin substitution can be calculated as $L_{new}$(mmol/g) = $L_{old}/(1 + L_{old}(W_{new} - W_{old}) \times 10^{-3})$, where L is the loading (mmol of amine per gram of resin), and W is the weight ($gmol^{-1}$) of the growing polyamide attached to the resin. see: Barlos, et al. *Int. J. Peptide Protein Res.* **1991**, *37*, 513.) was treated a solution of Boc-anhydride (500 mg) and DIEA (1 mL) in DMF (4 ml) and heated (55 °C) with periodic agitation for 30 min. ImPyPy-(*R*)$^{Boc}$γ-PyPyPy-β-Pam-resin was isolated by filtration, and washed sequentially with an excess of DMF, DCM, MeOH, and ethyl ether and the dried *in vacuo*. A sample of ImPyPy-(*R*)$^{Boc}$γ-PyPyPy-β-Pam-resin (240 mg, 0.18 mmol/gram) was treated with neat 3,3'-diamino-*N*-methyldipropylamine (2 mL) and heated (55 °C) with periodic agitation for 16h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC to provide ImPyPy-(*R*)$^{Boc}$γ-PyPyPy-β-Dp-NH$_2$ as a white powder upon lyophilization of the appropriate fractions (18 mg, 36% recovery); $[\alpha]^{20}_D$ -30 (*c* 0.05, $H_2O$); UV ($H_2O$) $\lambda_{max}$ 240, 306 (50,000); $^1$H NMR (DMSO-$d_6$) δ10.59 (s, 1 H), 10.16 (s, 1 H), 10.04 (m, 2 H), 10.00 (s,

1 H), 9.4 (br s, 1 H), 8.31 (d, 1 H, $J$ = 7.8 Hz), 8.16 (m, 2 H), 8.10 (t, 1 H, $J$ = 6.2 Hz), 7.89 (t, 1 H, $J$ = 5.8 Hz), 7.49 (s, 1 H), 7.37 (d, 1 H, $J$ = 1.7 Hz), 7.22 (d, 1 H, $J$ = 1.6 Hz), 7.21 (d, 1 H, $J$ = 1.5 Hz), 7.16 (m, 2 H), 7.14 (d, 1 H, $J$ = 1.6 Hz), 7.03 (d, 1 H, $J$ = 1.7 Hz), 6.99 (d, 1 H, $J$ = 1.4 Hz), 6.95 (d, 1 H, $J$ = 1.6 Hz), 6.91 (d, 1 H, J = 1.5 Hz), 6.78 (d, 1 H, $J$ = 1.5 Hz), 5.33 (m, 1 H), 3.95 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.76 (s, 3 H), 3.37 (q, 2 H, $J$ = 6.2 Hz), 3.07 (q, 2 H, $J$ = 5.9 Hz), 2.29 (t, 2 H, $J$ = 7.1 Hz), 1.93 (q, 2 H, $J$ = 5.8 Hz), 1.20 (m, 4 H); MALDI-TOF-MS (monoisotopic), 1135.3 (1135.6 calc. for $C_{54}H_{75}N_{18}O_{10}$).

**J. ImPyPy-(*S*)$^{Boc}$γ-PyPyPy-β-Dp-NH$_2$ (4-*S*-Boc-NH$_2$)**

ImPyPy-(*S*)$^{Boc}$γ-PyPyPy-β-Dp-NH$_2$ was prepared as described for **4-*R***. (16 mg, 32% recovery). $[\alpha]^{20}_D$ -30 (*c* 0.05, $H_2O$); $^1$H NMR (DMSO-$d_6$) is identical to **4-*R*-Boc-NH$_2$**; MALDI-TOF-MS (monoisotopic), 1135.4 (1135.6 calc. for $C_{54}H_{75}N_{18}O_{10}$).

**K. ImPyPy-(*R*)$^{Bo}$γPyPyPy-β-Dp-EDTA (4-*R*-Boc)**

Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImPyPy-$^{Boc}$γ-PyPyPy-β-Dp-NH$_2$ (10.4 mg, 10 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min.) and the remaining EDTA-anhydride hydrolyzed (0.1M NaOH, 3 mL, 55 °C, 10 min). Aqueous TFA (0.1% wt/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide ImPyPy-(*R*)$^{Boc}$γ-PyPyPy-β-Dp-EDTA (**4-*R*-Boc**) as a white powder upon lyophilization of the appropriate fractions (4 mg, 40% recovery). MALDI-TOF-MS (monoisotopic), 1409.6 (1409.7 calc. for $C_{64}H_{89}N_{20}O_{17}$).

**L. ImPyPy-(*S*)$^{Boc}$γ-PyPyPy-β-Dp-EDTA (4-*S*-Boc)**

ImPyPy-(*S*)$^{Boc}$γPyPyPy-β-Dp-NH$_2$ (12.0 mg, 12 μmol) was converted to **4-*S*-Boc** as described for **4-*R*-Boc** (4 mg, 33% recovery). MALDI-TOF-MS (monoisotopic), 1409.7 (1409.7 calc. for $C_{64}H_{89}N_{20}O_{17}$).

**M. ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Dp-EDTA (4-*R*)**

A sample of ImPyPy-$(R)^{Boc}$γ-PyPyPy-β-Dp-EDTA (2.1 mg) in DMSO (750 μL) was placed in a 50 mL flask and treated with TFA (15 mL, 22 °C, 2 h). Excess TFA was removed *in vacuo*, water added (6 mL) and the resulting solution purified by reversed phase HPLC to provide ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp-EDTA as a white powder upon lyophilization of the appropriate fractions (1.3 mg, 50% recovery). MALDI-TOF-MS (monoisotopic), 1309.5 (1309.6 calc. for $C_{59}H_{81}N_{20}O_{15}$).

**N. ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp-EDTA (4-*S*)**

ImPyPy-$(S)^{Boc}$γ–PyPyPy-β-Dp-EDTA (3.0 mg) was converted to **4-*S*** as described for **4-*R*** (1 mg, 33% recovery). MALDI-TOF-MS (monoisotopic), 1309.5 (1309.6 calc. for $C_{59}H_{81}N_{20}O_{15}$).

**O. ImPyPy-$(R)^{EDTA}$γ-PyPyPy-β-Dp (5-*R*)**

Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-Dp (1.0 mg, 1 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min.) and remaining EDTA-anhydride was hydrolyzed (0.1M NaOH, 3mL, 55 °C, 10 min.). Aqueous TFA (0.1% wt/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide **5-*R*** as a white powder upon lyophilization of the appropriate fractions (0.6 mg, 60% recovery). MALDI-TOF-MS (monoisotopic), 1266.4 (1266.6 calc. for $C_{57}H_{76}N_{19}O_{15}$). The *5-R* polyamide is shown in Figure 3.

**P. ImPyPy-$(S)^{EΔTA}$γ–PyPyPy-β-Dp (5-*S*)**

ImPyPy-$(S)^{EDTA}$γ-PyPyPy-β-Dp was prepared from **1-*S*** as described for **5-*R*** (6.8 mg, 16% recovery). MALDI-TOF-MS (monoisotopic), 1266.5 (1266.6 calc. for $C_{57}H_{76}N_{19}O_{15}$).

## Example 2

***Binding Site Size and Location by MPE•Fe(II) Footprinting***

**A. Preparation of 3′- and 5′-End-Labeled Restriction Fragments**

The plasmid pMM5 was linearized with *Eco*RI and *Bsr*BI, then treated with the Sequenase enzyme, deoxyadenosine 5′-[γ-$^{32}$P]triphosphate and thymidine 5′-[γ$^{32}$P]triphosphate for 3′ labeling. Alternatively, pMM5 was linearized with *Eco*RI, treated with calf alkaline phosphatase, and then 5′ labeled with T4 polynucleotide kinase and deoxyadenosine 5′-[γ$^{32}$P]triphosphate. The 5′ labeled fragment was then digested with *Bsr*BI. The labeled fragment (3′ or 5′) was loaded onto a 6% non-denaturing polyacrylamide gel, and the desired 135 base pair band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published methods (Iverson, et al. *Nucl. Acids Res.* **1987**, *15*, 7823; Maxam, et al. *Methods Enzymol.* **1980**, *65*, 499).

**B. MPE•Fe(II) Footprinting**

All reactions were carried out in a volume of 40 µL. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 25 mM Tris-acetate buffer (pH 7.0), 10 mM NaCl, 100 µM/base pair calf thymus DNA, and 30 kcpm 3′- or 5′-radiolabeled DNA. The solutions were allowed to equilibrate for 4 hours. A fresh 50 µM MPE•Fe(II) solution was prepared from 100 µL of a 100 µM MPE solution and 100 µL of a 100 µM ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$) solution. MPE•Fe(II) solution (5 µM) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 5 minutes. Cleavage was initiated by the addition of dithiothreitol (5 mM) and allowed to proceed for 14 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and a 5 µL sample (~ 15 kcpm) was immediately loaded onto an 8% denaturing polyacrylamide gel (5% crosslink, 7 M urea) at 2000 V.

**C. Results**

MPE•Fe(II) footprinting (Van Dyke, et al. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5470; Van Dyke, et al. *Science* **1984**, *225*, 1122) on 3′- and 5′-$^{32}$P end-labeled 135 base

pair restriction fragments reveals that the polyamides, each at 1 μM concentration, bind to the 5′-TGTTA-3′ match site (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair calf thymus DNA, pH 7.0 and 22 °C) (Figure 5 and 6). Compounds **1-*R*** and **3-*R*** each at 1.25 M, protect both the cognate 5'-TGTAA-3' site and the single base pair mismatch sequence 5′-TGTCA-3′. Remarkably, binding sequence preferences vary for the polyamides depending on the stereochemistry of the amine substituent. At 1.25 μM and 2.5 M concentration respectively, polyamides **1-*S*** and **3-*S*** bind a 5'-ACATT-3' reverse orientation match site in addition to the target match site 5'-TGTTA-3'. The sizes of the asymmetrically 3'-shifted footprint cleavage protection patterns for the polyamides are consistant with 5 base pair binding sites.

## Example 3

### *Binding Orientation Determination by Affinity Cleaving*

Affinity cleavage experiments (Taylor, et al. *Tetrahedron* **1984**, *40*, 457; Dervan, P. B. *Science* **1986**, *232*, 464) using hairpin polyamides modified with EDTA•Fe(II) at either the C-terminus or on the γ-turn, were used to determine polyamide binding orientation and stoichiometry. All reactions were carried out in a volume of 40 μL. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 25 mM Tris-acetate buffer (pH 7.0), 20 mM NaCl, 100 μM/base pair calf thymus DNA, and 20 kcpm 3′- or 5′-radiolabeled DNA. The solutions were allowed to equilibrate for 8 hours. A fresh solution of ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$) (10 μM) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 15 minutes. Cleavage was initiated by the addition of dithiothreitol (10 mM) and allowed to proceed for 30 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and the entire sample was immediately loaded onto an 8% denaturing polyacrylamide gel (5% crosslink, 7 M urea) at 2000 V.

Affinity cleavage experiments were performed on the same 3'- and 5'-$^{32}$P end-labeled 135 base pair restriction fragment (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair calf thymus DNA, pH 7.0 and 22 °C). The observed cleavage patterns for ImPyPy-(*R*)$^{H2N}$γ-PyPyPy-β-Dp-EDTA•Fe(II) (**4-*R*•Fe(II)**), ImPyPy-(*R*)$^{EDTA•Fe(II)}$γ–PyPyPy-β-Dp (**5-*R*•Fe(II)**), ImPyPy-(*S*)$^{H2N}$γ-PyPyPy-β-Dp-EDTA•Fe(II) (**4-*S*•Fe(II)**), ImPyPy-(*S*)$^{EDTA•Fe(II)}$γ-PyPyPy-β-Dp (**5-*S*•Fe(II)**) (Figures 7, 8 and 9) are in all cases 3'-shifted, consistent with minor groove occupancy. In the presence of 3.3 μM of **4-*R*•Fe(II)** and 10 μM **4-*S*•Fe(II)** which have an EDTA•Fe(II) moiety at the C-terminus, a single cleavage locus proximal to the 5' side of the 5'-TGTTA-3' match sequence is revealed. In the presence of 3.3 μM **5-*R*•Fe(II)** and 10 μM **5-*S*•Fe(II)** which have an EDTA•Fe(II) moiety appended to the γ-turn, a single cleavage locus is revealed proximal to the 3' side of the 5'-TGTTA-3' match sequence. Cleavage loci are more concise for the γ-turn EDTA•Fe(II) placement relative to carboxy terminal placement, consistent with the shorter tether. Cleavage loci are observed at both the 5' and 3' side of the 5'-TGTCA-3' single base pair mismatch site in the presence of 10 μM of **4-*R*•Fe(II).** The cleavage patterns observed at the 3' side of the site is approximately 3-fold more intense than cleavage at the 5'-side. For polyamide **4-*S*•Fe(II)** at 10 μM concentration, a single cleavage locus is revealed proximal to the 5' side of the 5'-ACATT-3' reverse orientation match site.

## Example 4

### *Energetics by Quantitative DNase I Footprinting*

### A. DNase I Footprinting

All reactions were carried out in a volume of 400 μL. Carrier DNA was not used in these reactions until after DNase I cleavage. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 10 mM Tris•HCl buffer (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, and 30 kcpm 3'-radiolabeled DNA. The solutions were allowed to equilibrate for a minimum of 12

hours at 22 °C. Cleavage was initiated by the addition of 10 μL of a DNase I stock solution (diluted with 1 mM DTT to give a stock concentration of 1.875 u/mL) and was allowed to proceed for 7 min at 22 °C. The reactions were stopped by adding 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 0.6 mg/mL glycogen, and 30 μM base-pair calf thymus DNA, and then ethanol precipitated. The cleavage products were resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% crosslink, 7 M urea) at 2000 V for 1 hour. The gels were dried under vacuum at 80 °C, then quantitated using storage phosphor technology.

Equilibrium association constants were determined as previously described (Mrksich, et al. *J. Am. Chem. Soc.* **1994**, *116*, 7983). The data were analyzed by performing volume integrations of the 5′-TGTTA-3′ and 5′-TGACA-3 sites and a reference site. The apparent DNA target site saturation, $\theta_{app}$, was calculated for each concentration of polyamide using the following equation:

$$\theta_{app} = 1 - \frac{I_{tot}/I_{ref}}{I_{tot}^{\circ}/I_{ref}^{\circ}} \qquad (1)$$

where $I_{tot}$ and $I_{ref}$ are the integrated volumes of the target and reference sites, respectively, and $I_{tot}^{\circ}$ and $I_{ref}^{\circ}$ correspond to those values for a DNase I control lane to which no polyamide has been added. The ($[L]_{tot}$, $\theta_{app}$) data points were fit to a Langmuir binding isotherm (eq 2, n=1 for polyamides **1-3**, n=2 for polyamides **4** and **5**) by minimizing the difference between $\theta_{app}$ and $\theta_{fit}$, using the modified Hill equation:

$$\theta_{fit} = \theta_{min} + (\theta_{max} - \theta_{min}) \frac{K_a^{n}[L]^{n}_{tot}}{1 + K_a^{n}[L]^{n}_{tot}} \qquad (2)$$

where $[L]_{tot}$ corresponds to the total polyamide concentration, $K_a$ corresponds to the equilibrium association constant, and $\theta_{min}$ and $\theta_{max}$ represent the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. Data were fit using a nonlinear least-squares fitting procedure of KaleidaGraph software (version 2.1, Abelbeck software) with $K_a$, $\theta_{max}$, and $\theta_{min}$ as the adjustable parameters. All acceptable fits had a correlation coefficient of $R > 0.97$. At least three sets of acceptable data were used in determining each association constant. All lanes from each gel were used unless visual inspection revealed a data point to be obviously flawed relative to neighboring points. The data were normalized using the following equation:

$$\theta_{norm} = \frac{\theta_{app} - \theta_{min}}{\theta_{max} - \theta_{min}} \qquad (3)$$

Photostimulable storage phosphorimaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark at 22 °C for 12-20 h. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integrations of all bands using the ImageQuant v. 3.2.

### B. Results

Quantitative DNase I footprint titrations (Brenowitz, et al.. *Methods Enzymol.* 1986, *130*, 132; Brenowitz, et al. *Proc. Natl. Acad. Sci. U.S.A.* 1986, *83*, 8462; Senear, et al. *Biochemistry* 1986, *25*, 7344) (10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$, pH 7.0 and 22 °C) were performed to determine the equilibrium association constant ($K_a$) of each six-ring hairpin polyamide for the three resolved sites (Figure 10 and 11). The 5′-TGTTA-3′ site is bound by polyamides in the order: ImPyPy-(*R*)$^{H2N}$γ-PyPyPy-β-Dp (1-*R*) ($K_a = 3.8 \times 10^9\ M^{-1}$) _ ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-EtOH (2-*R*) ($K_a = 3.3 \times 10^9\ M^{-1}$) > ImPyPy-(*R*)$^{Ac}$γ-PyPyPy-β-Dp (3-*R*) ($K_a = 3.0 \times 10^8\ M^{-1}$) _ ImPyPy-γ-PyPyPy-β-Dp ($K_a = 2.9 \times 10^8\ M^{-1}$) > ImPyPy-(*S*)$^{H_2N}$γPyPyPy-β-Dp (1-*S*) ($K_a = 2.2 \times 10^7$

$M^{-1}$) > ImPyPy-(*S*)$^{Ac}$γ-PyPyPy-β-Dp (3-*S*) ($K_a < 5.0 \times 10^6\ M^{-1}$). Equilibrium association constants for recognition of the 5'-TGACT-3' single base pair mismatch site are: ImPyPy-(*R*)$^{H_2N}$γ–PyPyPy-β-Dp (1-*R*) ($K_a = 3.5 \times 10^7\ M^{-1}$) _ ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-EtOH (2-*R*) ($K_a = 3.1 \times 10^7\ M^{-1}$) > ImPyPy-(*R*)$^{Ac}$γ-PyPyPy-β-Dp (3-*R*) ($K_a < 5 \times 10^6\ M^{-1}$) _ ImPyPy-γ-PyPyPy-β-Dp ($K_a = 4.8 \times 10^6\ M^{-1}$). The polyamides ImPyPy-(*S*)$^{H_2N}$γ-PyPyPy-β-Dp (1-*S*) and ImPyPy-(*S*)$^{A\chi}$γPyPyPy-β-Dp (3-*S*) recognize the 5'-ACATT-3' reverse orientation sequence with $K_a = 4.6 \times 10^6\ M^{-1}$ and $K_a < 5 \times 10^6\ M^{-1}$ respectively. It should be noted that a detailed comparison of the relative mismatch binding energetics cannot be made since the 5'-TGACA-3' and 5'-ACATT-3' binding sites overlap. The relative affinity of 5'-TGTTA-3' match site binding varies from 100-fold to 5-fold depending on the stereochemistry of the γ-turn substitutions (Table 2).

**TABLE 2**

*Equilibrium Association Constants ($M^{-1}$)*[a,b]

| Improved Polymide | Match Site 5'-TGTTA-3' | Reverse Site 5'-ACATT-3' | Mismatch Site 5'-TGACA-3' | Specificity[c] |
|---|---|---|---|---|
| ImPyPy-γ-PyPyPy-β-Dp | $2.9 \times 10^8$ | ND | $4.8 \times 10^6$ | 60 |
| ImPyPy-(R)$^{H2N}$γ-PyPyPy-β-Dp | $3 \times 10^9$ (0.2) | ND | $3.5 \times 10^7$ (1.0) | 100 |
| ImPyPy-(S)$^{H2N}$γ-PyPyPy-β-Dp | $2.2 \times 10^7$ (0.7) | $4.6 \times 10^6$ (2.0)[d] | ND | 5 |
| ImPyPy-(R)$^{H2N}$γ-PyPyPy-β-EtOH | $3.3 \times 10^9$ (0.9) | ND | $3.1 \times 10^7$ (0.4) | 100 |
| ImPyPy-(R)$^{Ac}$γ-PyPyPy-β-Dp | $3.0 \times 10^8$ (1.3) | ND | $<5.0 \times 10^6$ | ≥60 |
| ImPyPy-(S)$^{Ac}$γ-PyPyPy-β-Dp | $<5.0 \times 10^6$ | $<5.0 \times 10^6$ | ND | ND |

[a]The reported association constants are the average values obtained from three DNase I footprint titration experiments. The standard deviation for each data set is indicated in parentheses. The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

[b]The five base piar binding sites are in captital letters.

[c]Specificity is calculated by $K_a$ (match site)/$K_a$ (mismatch site).

[d]Mismatch site is 5'(ACATT)-3' for ImPyPy-(S)$^{H2N}$γ-PyPyPy-β-Dp_(1-S) and ImPyPy-(S)$^{Ac}$γ-PyPyPy-β-Dp (3-S) as determined by MPE•Fe(II) footprinting and affinity cleaving.

ND = not determined.

## Example 5

### *Binding site size and orientation*

MPE•Fe(II) footprinting reveals that the polyamides bind with highest affinity to the 5'-TGTTA-3' match site, the 5'-TGACA-3' single base pair mismatch site for polyamides **1-*R*** and **3-*R***, and the 5'-ACATT-3' reverse orientation match site for polyamides **1-*S*** and **3-*S*** (Figure 6). Affinity cleaving experiments using polyamides with EDTA•Fe(II) placed at either the carboxy terminus or the γ-turn confirm that polyamides derived from both (*R*) and (*S*)-2,4-diaminobutyric acid bind to the 5'-TGTTA-3' target site with a single orientation (Figure 10). The observation of a single cleavage locus is consistent only with an oriented 1:1 polyamide:DNA complex in the minor groove and rules out any dimeric overlapped or extended binding motifs. The hairpin binding model is further supported by the location of the cleavage locus at either the 5' or 3' side of the 5'-TGTTA-3' target site corresponding to EDTA•Fe(II) placement at the polyamide carboxy terminus or the γ-turn, respectively (Figure 10). Polyamide subunits linked by the $(R)^{H_2N}\gamma\gamma$ bind the symmetric single base pair mismatch sequence 5'-TGACA-3' in two distinct orientations. Polyamides linked with $(S)^{H_2N}\gamma$ bind to a 5'-ACATT-3' reverse orientation match sequence as revealed by a unique cleavage locus at the 5' side of the site.

## Example 6

### *Binding Affinity*

All six polyamides bind to the 5'-TGTTA-3' target site with binding isotherms (eq. 2, n = 1) consistent with binding as an intramolecular hairpin (Figure 11). However the relative match site binding affinity varies by nearly 1000-fold depending on the stereochemistry of the γ-turn and the nature of the substituents. Among the six polyamides, ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1-*R***) binds to the targeted 5'-TGTTA-3' site with the highest affinity. ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp binds with an equilibrium association constant, ($K_a = 3 \times 10^9\ M^{-1}$; Parks, et al. *J. Am. Chem. Soc.* **1996**, *118*, 6147),

a factor of 10 greater than that of the parent polyamide, ImPyPy-γ-PyPyPy-β-Dp, ($K_a$ = 3 x $10^8$ $M^{-1}$). Replacement of the C-terminal dimethylaminopropylamide group of **1-*R*** with an ethoxyamide group as in ImPyPy-$(R)^{H_2N}$γ-PyPyPy-β-EtOH (**2-*R***) results in no decrease in binding affinity ($K_a$ = 3 x $10^9$ $M^{-1}$). Acetylation of the γ-turn amino group as in ImPyPy-$(R)^{Ac}$γ-PyPyPy-β-Dp (**3-*R***) reduces binding affinity 10-fold ($K_a$ = 3 x $10^8$ $M^{-1}$) relative to **1-*R***.

The observation that polyamides which differ only by replacement of the dimethylaminopropylamide group **1-*R*** with an ethoxyamide group **2-*R*** bind with similar affinity indicates that interactions between the cationic dimethylaminopropyl tail group with anionic phosphate residues or the negative electrostatic potential in the floor of the minor groove may not contribute substantially to the energetics of hairpin-DNA binding (Zimmer, et al. *Prog. Biophys. Molec. Biol.* **1986**, *47*, 31; Pullman, B. *Adv. Drug. Res.* **1990**, *18*, 1; Breslauer, et al. *Structure and Expression (Vol. 2), DNA and Its Drug Complexes* p. 273-289, R. H. Sarma and M. H. Sarma (eds.) Academic Press (1988)). Furthermore, these results indicate that the observed binding enhancement of **1-*R***, in relation to ImPyPy-γ-PyPyPy-β-Dp, is not simply the difference between a monocationic and dicationic ligand binding to the polycationic DNA helix (Zimmer, et al. *Prog. Biophys. Molec. Biol.* **1986**, *47*, 31; Pullman, B. *Adv. Drug. Res.* **1990**, *18*, 1; Breslauer, et al. *Structure and Expression (Vol. 2), DNA and Its Drug Complexes* p. 273-289, R. H. Sarma and M. H. Sarma (eds.) Academic Press (1988)). The modest increased binding affinity of polyamide **1-*R*** may result from electrostatic interactions between the precisely placed amine group and the floor of the minor groove. Alternately, the increased affinity could indicate a reduction in the degrees of freedom accessible to the free hairpin in solution resulting from a steric effect, or an electrostatic interaction between the positively charged amine group and the negative potential of the γ–carbonyl group.

Polyamides linked with $(S)^{H_2N}$γ, ImPyPy-$(S)^{H_2N}$γ-PyPyPy-β-Dp (**1-*S***) and ImPyPy-$(S)^{Ac}$γ-PyPyPy-β-Dp (**3-*S***), bind to the 5'-TGTTA-3' match site with 100-fold

($K_a$ = 2 x $10^7$ $M^{-1}$) and 1000-fold ($K_a$ < 5 x $10^6$ $M^{-1}$) reduced affinity relative to the $(R)^{H_2N}\gamma$ linked polyamide **1-*R***. These results demonstrate that the DNA-binding affinity of chiral hairpin polyamides can be predictably regulated as a function of the stereochemistry of the turn residue.

## Example 7

### *Sequence-specificity*

Polyamides with a variety of substitutions at the γ-turn bind preferentially to the 5'-TGTTA-3' match site, while overall specificity versus binding at reverse orientation and mismatch sites is modified. Replacing the α-proton in the γ–residue of ImPyPy-γ-PyPyPy-β-Dp with an amino group that confers a chiral α-hydrogen (*R*) configuration, provides the most specific polyamide ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1-*R***). The ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp•5'-TGTTA-3' complex forms with 100-fold preference relative to the ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp•5'-TGTCA-3' mismatch complex. Substitution of the charged dimethylaminopropyl tail group with an ethoxyamide group as in (**2-*R***) does not alter binding specificity. The modest increase in specificity against single base mismatch sequences for polyamides **1-*R*** and **2-*R*** (100-fold) relative to the parent unsubstituted hairpin polyamide (60-fold) implicates chiral hairpin polyamides as an optimized class of small molecules for recognition of the DNA minor groove.

## Example 8

### *Binding Orientation*

In principle, a polyamide:DNA complex can form at two different DNA sequences depending on the alignment of the polyamide (N-C) with the walls of the minor groove (5'-3'; White, et al. *J. Am. Chem. Soc.* **1997,** *119*, 8756). A six ring-hairpin polyamide of core sequence composition ImPyPy-γ-PyPyPy which places the N-terminus of each three-ring polyamide subunit at the 5'-side of individual recognized DNA strands would bind to 'forward match' 5'-WGWWW-3' sequences (W = A or T). Placement of the polyamide N-terminus at the 3'-side of each recognized strand would result in

targeting 'reverse match' 5'-WCWWW-3' sequences. For hairpin polyamides there is an energetic preference for 'forward' alignment of each polyamide subunit (N-C) with respect to the backbone (5'-3') of the DNA double helix (White, et al. *J. Am. Chem. Soc.* **1997,** *119,* 8756).

In addition to decreasing the affinity for the 5'-TGTTA-3' match site, replacing the α-proton of γ-turn in ImPyPy-γ–PyPyPy-β-Dp with $(S)^{H_2N}\gamma$ changes the mismatch sequence preference from the 5'-TGACA-3' site bound by the $(R)^{H_2N}\gamma$-linked polyamides to a 5'-ACATT-3' reverse match site. Binding to the reverse site may result from the presence of the steric bulk of the amino or acetamido groups in the floor of the minor groove preventing the deep polyamide binding required for specific DNA recognition. However, an analysis of hairpin folding requirements for 'forward' and 'reverse' binding reveals an additional model.

In principle, there exist two non-superimposable hairpin folds which are related by mirror plane symmetry (Figure 12). One hairpin fold is responsible for the preferred 5' to 3' N to C orientation, while the other fold corresponds to the 3' to 5' N to C reverse orientation binding. For an achiral hairpin polyamide in the absence of DNA, each non-superimposable fold should be energetically equivalent. However, an asymmetrically folded hairpin polyamide with a chiral substituent could potentially display differential energetics for oriented binding (Figure 12). In the forward folded hairpin (5' to 3' N to C), $(R)^{H_2N}\gamma$ directs the amine functionality away from the DNA helix, while (*S*) enantiomer is predicted to direct the amine into the floor of the minor groove. For the 'reverse' fold hairpin, $(S)^{H_2N}\gamma$ directs the amine functionality away from the floor of the DNA helix, while the amine of the (*R*) enantiomer is predicted to clash with the floor of the helix. The modest enhanced specificity of chiral polyamides **1-*R*** and **2-*R*** relative to the unsubstituted parent hairpin may result from stabilization of the forward binding mode and/or destabilization of the reverse binding hairpin fold.

## Example 9

### Tandemly-Linked Polyamides

**A. Synthesis**

ImPyPy-(*R*)[ImPyPy-(*R*)$^{H_2N}\gamma$–PyPyPy-β-]$^{HN}\gamma$-PyPyPy-β-Dp (**6**) and ImPyPy-(*R*)[ImPyPy-(*R*)$^{H_2N}\gamma$–PyPyPy-δ-]$^{HN}\gamma$-PyPyPy-β-Dp (**7**) were synthesized from Boc-β-alanine-Pam resin (0.6 g resin, 0.6 mmol/g substitution) using Boc-chemistry machine-assisted protocols in 31 steps (Figure 3).[12] ImPyPy-(*R*)$^{FmocHN}\gamma$–PyPyPy-β-Pam-Pam resin was prepared as described.[10] The Fmoc protecting group was then removed by treatment with (4:1) piperidine/DMF. The remaining amino acid sequence was then synthesized in a stepwise manner using Boc-chemistry machine assisted protocols to provide ImPyPy-(*R*)[ImPyPy-(*R*)$^{FmocHN}\gamma$–PyPyPy-β-]$^{HN}\gamma$-PyPyPy-β-Pam-Resin and ImPyPy-(*R*)[ImPyPy-(*R*)$^{FmocHN}\gamma$–PyPyPy-δ-]$^{HN}\gamma$-PyPyPy-β-Pam-Resin. The Fmoc group was removed with (4:1) piperdine/DMF. A sample of resin was then cleaved by a single-step aminolysis reaction with ((dimethylamino)propylamine (55 °C, 18 h) and the reaction mixture subsequently purified by reversed phase HPLC to provide ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}\gamma$–PyPyPy-β-]$^{HN}\gamma$-PyPyPy-β-Dp (**6**) and ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}\gamma$–PyPyPy-δ-]$^{HN}\gamma$-PyPyPy-β-Dp (**7**). For the synthesis of the EDTA-turn derivative **7-E**, a sample ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}\gamma$–PyPyPy-δ-]$^{HN}\gamma$-PyPyPy-β-Dp (**7**) was treated with an excess of EDTA-dianhydride (DMSO/NMP, DIEA 55 °C, 30 min.) and the remaining anhydride hydrolyzed (0.1 M NaOH, 55 °C, 10 min.). The polyamide ImPyPy-(*R*)[ImPyPy-(R)$^{EDTA}\gamma$–PyPyPy-δ-]$^{HN}\gamma$-PyPyPy-β-Dp (**7-E**) was then isolated by reverse phase HPLC. The dicationic twelve-ring tandem hairpins are soluble at concentrations up to 1 mM. The solubility of the tandem hairpins is 10- to 100-fold greater than that found for extended or hairpin twelve-ring polyamides.

**ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}\gamma$–PyPyPy-β-]$^{HN}\gamma$-PyPyPy-β-Dp ("6")**

ImPyPy-(*R*)$^{Fmoc}\gamma$-PyPyPy-β-Pam-Resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-Pam-Resin (0.6 mmol/g). (*R*)-2-Fmoc-4-Boc-diaminobutyric acid (0.7 mmol) was incorporated as previously

described for Boc-γ-aminobutyric acid. ImPyPy-(*R*)$^{Fmoc}$γ-PyPyPy-β-Pam-Resin was placed in a glass 20 mL peptide synthesis vessel and treated with DMF (2 mL), followed by piperidine (8 mL) and agitated (22 °C, 30 min.). ImPyPy-(*R*)$^{H2N}$γ-PyPyPy-β-Pam-resin was isolated by filtration, and washed sequentially with an excess of DMF, DCM, MeOH, and ethyl ether and the amine-resin dried *in vacuo*. ImPyPy-(*R*)[ImPyPy-(R)$^{FMocHN}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Pam-Resin was then synthesized in a stepwise fashion by machine-assisted solid phase methods from ImPyPy-(*R*)$^{H2N}$γ-PyPyPy-β-Pam-resin (0.38 mmol/g). ImPyPy-(*R*)[ImPyPy-(R)$^{FMocHN}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Pam-Resin was placed in a glass 20 mL peptide synthesis vessel and treated with DMF (2 mL), followed by piperidine (8 mL) and agitated (22 °C, 30 min.). ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Pam-Resin was isolated by filtration, and washed sequentially with an excess of DMF, DCM, MeOH, and ethyl ether and the amine-resin dried *in vacuo*. A sample of ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Pam-Resin (240 mg, 0.29 mmol/gram) was treated with neat dimethylaminopropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (28 mg, 22% recovery). UV ($H_2O$) $\lambda_{max}$ 246, 306 (100,000); MALDI-TOF-MS [$M^+$–H] (monoisotopic), 1881.9: 1881.9 calc. for $C_{89}H_{109}N_{32}O_{16}$

**ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp ("7")**

ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Pam-Resin was prepared as described for ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Pam-Resin. A sample of ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Pam-Resin (240 mg, 0.29 mmol/gram[19]) was treated with neat dimethylaminopropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL) and the resulting solution purified by reversed phase HPLC. ImPyPy-

(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (32 mg, 25% recovery). $[\alpha]^{20}_{D}$ +14.6 (*c* 0.05, $H_2O$); UV ($H_2O$) $\lambda_{max}$ 246, 306 (100,000); $^1$H NMR (300 MHz, [$D_6$]DMSO, 20 °C): δ = 10.54 (s, 1 H; aromatic NH); 10.45 (s, 1 H; aromatic NH); 10.44 (s, 1 H; aromatic NH); 10.02 (s, 1 H; aromatic NH); 9.95 (s, 1 H; aromatic NH); 9.92 (s, 1 H; aromatic NH); 9.90 (d, 2 H; aromatic NH); 9.86 (d, 2 H; aromatic NH); 9.2 (br s, 1 H; $CF_3COOH$); 8.25 (m, 4 H; aliphatic NH, $NH_3$); 8.11 (d, 1 H; *J* = 8.5 Hz, aliphatic NH); 8.04 (m, 4H, aliphatic NH), 7.37 (s, 2 H; CH); 7.25 (m, 2 H; CH); 7.22 (d, 1 H; CH); 7.18 (m, 2 H; CH); 7.16 (m, 3 H; CH); 7.12 (m, 4 H; CH); 7.02 (m, 4 H; CH); 6.95 (d, 1 H; *J* = 1.6 Hz; CH); 6.91 (d, 1 H; *J* = 1.5 Hz; CH); 6.88 (d, 1 H, *J* = 1.3 Hz; CH); 6.85 (m, 3 H; CH); 5.32 (t, 1 H; aliphatic CH), 4.45 (m, 1 H, aliphatic CH), 3.96 (s, 6 H; $NCH_3$); 3.83 (s, 3 H; $NCH_3$); 3.80 (s, 18 H; $NCH_3$); 3.79 (s, 3 H; $NCH_3$); 3.76 (s, 3 H; $NCH_3$); 3.39 (m, 4 H; $CH_2$); 3.28 (m, 2 H; $CH_2$); 3.15 (m, 4 H; $CH_2$); 3.07 (m, 2 H; $CH_2$); 2.97 (m, 2 H; $CH_2$); 2.70 (d, 6 H; $N(CH_3)_2$); 2.32 (m, 2 H; $CH_2$); 1.93 (m, 2 H; $CH_2$); 1.71 (m, 2 H; $CH_2$); 1.47 (m, 2 H; $CH_2$); 1.20 (m, 4 H; $CH_2$); MALDI-TOF-MS [$M^+$–H] (monoisotopic), 1910.2: 1909.9 calc. for $C_{91}H_{113}N_{32}O_{16}$.

**ImPyPy-(*R*)[ImPyPy-(R)$^{EDTA}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp ("7-E")**

Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp (10 mg, 5 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min.) and the remaining EDTA-anhydride hydrolyzed (0.1M NaOH, 3 mL, 55 °C, 10 min). Aqueous TFA (0.1% wt/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide ImPyPy-(*R*)[ImPyPy-(R)$^{EDTA}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp (7-E) as a white powder upon lyophilization of the appropriate fractions (2 mg, 20% recovery). MALDI-TOF-MS [$M^+$–H] (monoisotopic), 2184.3: 2184.0 calc. for $C_{101}H_{127}N_{34}O_{23}$

**B. Plasmid and restriction fragment preparation**

The plasmids pDH10, pDH11, and pDH12 were constructed by hybridization of the inserts listed in Figure 18. Each hybridized insert was ligated individually into linearized pUC19 *Bam*HI/*Hind*III plasmid using T4 DNA ligase. The resultant constructs were used to transform Top10F′ OneShot competent cells from Invitrogen. Ampicillin-resistant white colonies were selected from 25 mL Luria-Bertani medium agar plates containing 50 μg/mL ampicillin and treated with XGAL and IPTG solutions. Large-scale plasmid purification was performed with Qiagen Maxi purification kits. Dideoxy sequencing was used to verify the presence of the desired insert. Concentration of the prepared plasmid was determined at 260 nm using the relationship of 1 OD unit = 50 μg/mL duplex DNA.

The plasmids pDH(11-12) were linearized with *Eco*RI and *Bsr*BI, then treated with the Sequenase enzyme, deoxyadenosine 5′-[α-$^{32}$P]triphosphate and thymidine 5′-[α-$^{32}$P]triphosphate for 3′ labeling. Alternatively, these plasmids were linearized with *Eco*RI, treated with calf alkaline phosphatase, and then 5′ labeled with T4 polynucleotide kinase and deoxyadenosine 5′-[γ-$^{32}$P]triphosphate. The 5′ labeled fragment was then digested with *Bsr*BI. The labeled fragment (3′ or 5′) was loaded onto a 6% non-denaturing polyacrylamide gel, and the desired 147 base pair band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published methods.

**C. Binding Site Size**

All reactions were carried out in a volume of 40 μL. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 25 mM Tris-acetate buffer (pH 7.0), 10 mM NaCl, 100 μM/base pair calf thymus DNA, and 30 kcpm 3′- or 5′-radiolabeled DNA. The solutions were allowed to equilibrate for 4 hours. A fresh 50 μM MPE•Fe(II) solution was prepared from 100 μL of a 100 μM MPE solution and 100 μL of a 100 μM ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$) solution. MPE•Fe(II)

solution (5 μM) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 5 minutes. Cleavage was initiated by the addition of dithiothreitol (5 mM) and allowed to proceed for 14 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and a 5 μL sample (~ 15 kcpm) was immediately loaded onto an 8% denaturing polyacrylamide gel (5% crosslink, 7 M urea) at 2000 V.

MPE•Fe(II) footprinting on 3′- or 5′-$^{32}$P end-labeled 135 base pair *Eco*RI/*Bsr*BI restriction fragments from the plasmid pDH11 reveals that polyamide 7, at 100 pM concentration, binds to the designated 11-bp match site 5′-TGTTATTGTTA-3′ (25 mM Tris-acetate, 10 mM NaCl, pH 7.0 and 22 °C) (Figures 5a and 5c). Binding of the mismatch site 5′-TGTCATTGTCA-3′ is only observed at much higher polyamide concentrations (Figure 5a). The size of the asymmetrically 3′-shifted cleavage protection pattern for polyamide 7 at the designated match site 5′-TGTTATTGTTA-3′ is consistent with formation of the predicted hairpin-δ-hairpin•DNA complex.

**D. Binding Orientation**

All reactions were carried out in a volume of 40 μL. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were: 25 mM Tris-acetate buffer (pH 7.0), 20 mM NaCl, 100 μM/base pair calf thymus DNA, and 20 kcpm 3′- or 5′-radiolabeled DNA. The solutions were allowed to equilibrate for 8 hours. A fresh solution of ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2•6H_2O$) (10 μM) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 15 minutes. Cleavage was initiated by the addition of dithiothreitol (10 mM) and allowed to proceed for 30 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 6 min, and the entire sample was immediately loaded onto an 8% denaturing polyacrylamide gel (5% crosslink, 7 M urea) at 2000 V.

Affinity cleavage experiments using 7-E which has an EDTA•Fe(II) moiety appended to the γ-turn, were used to confirm polyamide binding orientation and

stoichiometry. Affinity cleavage experiments were performed on the same 3'- or 5'-$^{32}$P end-labeled 135 base pair DNA restriction fragment from the plasmid pDH11 (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair calf thymus DNA, pH 7.0 and 22 °C). The observed cleavage pattern for 7-E (Figures 5b and 5d) are 3'-shifted, consistent with minor groove occupancy. In the presence of 1 μM 7-E, a single cleavage locus proximal to the 3' side of the 5'-TGTTATTGTTA-3' match sequence is revealed, consistent with formation of an oriented 1:1 hairpin-δ-hairpin•DNA complex.

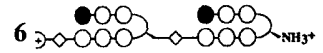
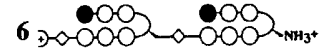
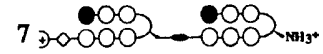
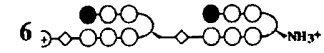
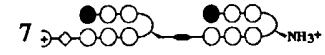
**E. Equilibrium Association Constants**

Quantitative DNase I footprinting and related mathematical calculations were performed as described above in Example 4, except as otherwise indicated below. Quantitative DNase I footprint titrations (10 mM Tris•HCl, 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$, pH 7.0 and 22 °C) were performed to determine the equilibrium association constants ($K_a$) of 6 and 7 for the 10-, 11- and 12-bp match and mismatch sites (Table 1). Polyamide 7 preferentially binds the 11-bp 5'-TGTTATTGTTA-3' target sequence with an equilibrium association constant, $K_a > 1 \times 10^{11}$ $M^{-1}$. The corresponding 11 bp mismatch 5'-TGTCATTGTCA-3' site is bound by 7 with > 4500-fold lower affinity ($K_a = 2.2 \times 10^8$ $M^{-1}$). Polyamide 7 binds the 10 bp site 5'-TGTTATGTTA-3' ($K_a = 1.5 \times 10^{10}$ $M^{-1}$) and the 12 bp site 5'-TGTTATATGTTA-3' ($K_a = 1.0 \times 10^9$ $M^{-1}$) with 70- and 1000-fold lower affinity, respectively. Polyamide 6 binds the 10-bp 5'-TGTTATGTTA-3' site and 11-bp 5'-TGTTATTGTTA-3' site with $K_a = 2 \times 10^{10}$ $M^{-1}$, and also binds the 12-bp 5'-TGTTATATGTTA-3' site with 16-fold lower affinity ($K_a = 9.0 \times 10^9$ $M^{-1}$). The parent hairpin ImPyPy-(*R*)$^{H_2N}$γ–PyPyPy-β-Dp was found to bind to the 5'-TGTTA-3' match site with $K_a = 5 \times 10^9$ $M^{-1}$.

Formally the subunits of polyamide 7 represent the combination of the parent-acetylated parent tandem hairpin. The parent and acetylated hairpins occupy I with binding energetics of -13.2 kcal/mol and -11.8 kcal/mol respectively, predicting that covalent linkage of parent-δ-acetylated would bind the same site

with an association constant of ($K_a$) = 2.2 x $10^{18}$ $M^{-1}$, 6 orders of magnitude higher than the observed tandem hairpin 7 which binds with -16.3 kcal/mol. Relative to parent recognition of the 5'-TGTTA-3' half site, we only observe a 2-fold enhancement and 5-fold decrease in binding respectively, recognition of the 10- and 12-bp sites. At site IV, a single base pair mismatch reduces binding eneretics for both unlinked and linked hairpins. Unlinked parent and acetylated-parent respectively bind with 10.5 kcal/mol and 9.2 kcal/mol, predicting the linked polyamide would bind with a ($K_a$) = 2.4 x $10^{14}$ $M^{-1}$. The observed tandem hairpin binds with energetics of -11.35 kcal/mol and a ($K_a$) = 2.2 x $10^8$ $M^{-1}$.

**Table 2** Equilibrium Association Constants ($M^{-1}$)[a,b]

| Polyamide | 5'-aTGTTATGTTAg-3' | 5'-aTGTCATGTCAt-3' | Specificity |
|---|---|---|---|
| 6  | **2 x $10^{10}$** | 1.5 x $10^8$ | 133 |
| 7 | **1.5 x $10^{10}$** | 1.9 x $10^8$ | 80 |
| **Polyamide** | **5'-aTGTTATTGTTAg-3'** | **5'-aTGTCATTGTCAt-3'** | **Specificity** |
| 6  | **1.5 x $10^{10}$** | 2 x $10^8$ | 75 |
| 7  | **≥ 1 x $10^{12}$** | 2.2 x $10^8$ | ≥ 4500 |
| **Polyamide** | **5'-aTGTTATATGTTAg-3'** | **5'-aTGTCATATGTCAt-3'** | **Specificity** |
| 6  | **9 x $10^8$** | 3 x $10^7$ | 30 |
| 7  | **1 x $10^9$** | 2.5 x $10^7$ | 40 |

[a]The reported association constants are the average values obtained from three DNase I footprint titration experiments. The assays were carried out at 22 °C at pH 7.0 in the presence of 10 mM Tris-HCL, 10 mM KCl, 10 mM $MgCl_2$, ad 5 mM $CaCl_2$. [b]The ten, eleven, and twelve base-pair sites are in capital letters, with flanking sequences in lower case letters. Match site association constants are shown in boldtype. Specificity is calculated as $K_a$(match)/$K_a$(mismatch).

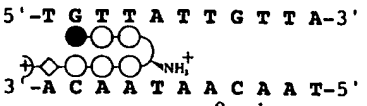
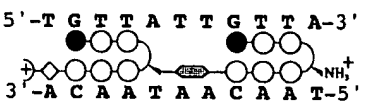
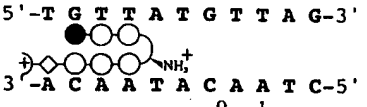
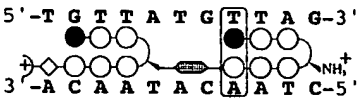
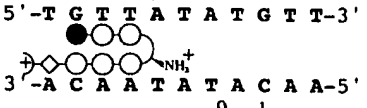
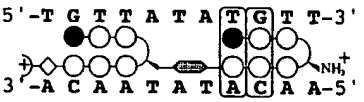
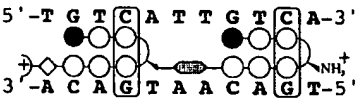
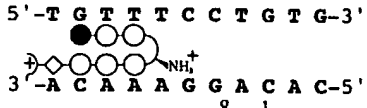
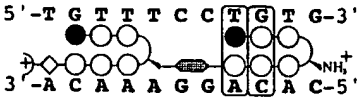
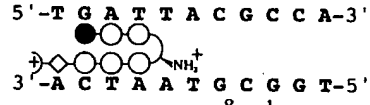
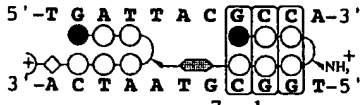
### F. Linker Dependence

Site size preferences of polyamides 6 and 7 result from modifications to the length if the turn-to-tail linker. Modeling indicated that β and δ linkers would provide sufficient length for recognition of either 10- or 11- base pairs, but would be too short to span the 12-bp binding site. Polyamide 7 displays optimal recognition of the 11-bp site, binding the 5'-TGTTATTGTTA-3' site with a $K_a$ _ 1 x $10^{12}$ $M^{-1}$.

Replacing the δ linker in 7 with the 2-carbon shorter β-alanine residue in 6 shows a reduction of affinity at the 11-bp site by > 6-fold ($K_a$ = 1.5 x $10^{10}$ $M^{-1}$). The highly reduced affinities of 6 and 7 at the 12-bp site indicates that covalent constaints of the linker subunit prevents alignment of hairpin subunits for their binding sites.

**Table 3**

***Binding affinity of 6-ring hairpin and δ-linked tandem hairpins at 11-bp match site I, and a series of mismatch sites II-VI***

| | Parent | Polyamide 2 |
|---|---|---|
| I |  $K_a$= 5 x $10^9$ $M^{-1}$ |  $K_a \geq$ 1 x $10^{12}$ $M^{-1}$ |
| II |  $K_a$= 5 x $10^9$ $M^{-1}$ |  $K_a$= 1.5 x $10^{10}$ $M^{-1}$ |
| III |  $K_a$= 5 x $10^9$ $M^{-1}$ |  $K_a$= 1 x $10^9$ $M^{-1}$ |
| IV | $K_a$= 5 x $10^7$ $M^{-1}$ |  $K_a$= 2.2 x $10^8$ $M^{-1}$ |
| V |  $K_a$= 3 x $10^8$ $M^{-1}$ |  $K_a$= 1 x $10^8$ $M^{-1}$ |
| VI |  $K_a$= 1 x $10^8$ $M^{-1}$ |  $K_a$= 5 x $10^7$ $M^{-1}$ |

As disclosed herein, the present invention provides the reagents and methodologies for the preparation and use of a variety of new chiral hairpin polyamide structures for specific recognition in the DNA minor groove. While a preferred form of the invention has been shown in the drawings and described, since variations in the preferred form will be apparent to those skilled in the art, the invention should not be construed as limited to the specific form shown and described, but instead is as set forth in the claims.

## CLAIMS

We claim:

1. In a polyamide having a hairpin turn derived from γ-aminobutyric acid which specifically binds to base pairs in the minor groove of a DNA molecule, the improvement comprising substitution of the γ-aminobutyric acid residue of the hairpin with (R)-2,4-diaminobutyric acid.

2. A polyamide of claim 1 wherein the R-2-amino is derivatized to form an acid amide.

3. A polyamide of claim 1 wherein the polyamide has three or four carboxyamide binding pairs.

4. A polyamide of claim 1 having the formula:

$$X_1X_2X_3\gamma X_4X_5X_6$$

wherein γ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_2$, $X_3/X_4$, and $X_5/X_6$ represent three carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound.

5. A polyamide of claim 1 having the formula:

$$X_1X_2X_3X_4\gamma X_5X_6X_7X_8$$

wherein γ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_8$, $X_2/X_7$, $X_3/X_6$, and $X_4/X_5$ represent four carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound.

6. A polyamide of claim 1 having the formula:

$$X_1X_2X_3X_4X_5\gamma X_6X_7X_8X_9X_{10}$$

wherein $\gamma$ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from $\gamma$-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_{10}$, $X_2/X_9$, $X_3/X_8$, $X_4/X_7$, $X_5/X_6$ represent five carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound.

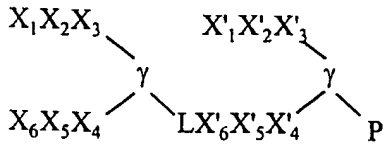
7. A polyamide of claim 1 having the formula:

$$X_1X_2X_3X_4X_5X_6\gamma X_7X_8X_9X_{10}X_{11}X_{12}$$

wherein $\gamma$ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from $\gamma$-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_{12}$, $X_2/X_{11}$, $X_3/X_{10}$, $X_4/X_9$, $X_5/X_8$, $X_6/X_7$ represent three or four carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound.

8. A tandem-linked polyamide of claim 1 having the formula:



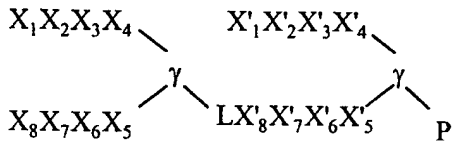
wherein γ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_6$, $X_2/X_5$, $X_3/X_4$, $X'_1/X'_6$, $X'_2/X'_5$, $X'_3/X'_4$ represent carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound;

L represents an amino acid linking group selected from the group consisting of β-alanine and 5-aminovaleric acid (δ);

P represents zero to ten improved polyamides of claim 1.

9. A tandem-linked polyamide of claim 1 having the formula:



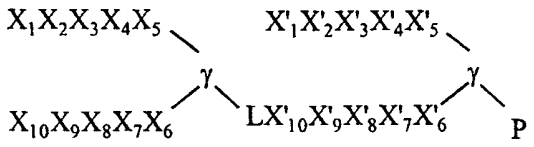
wherein γ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_8$, $X_2/X_7$, $X_3/X_6$, $X_4/X_5$, $X'_1/X'_8$, $X'_2/X'_7$, $X'_3/X'_6$, $X'_4/X'_5$ represent carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound;

L represents an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof; and,

P represents zero to ten improved polyamides of claim 1.

10. A tandem-linked polyamide of claim 1 having the formula:



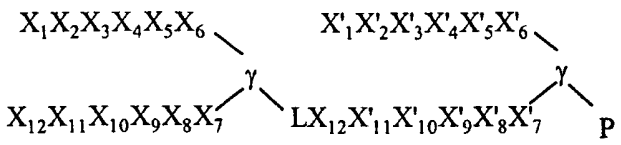
wherein γ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_{10}$, $X_2/X_9$, $X_3/X_8$, $X_4/X_7$, $X_5/X_6$, $X'_1/X'_{10}$, $X'_2/X'_9$, $X'_3/X'_8$, $X'_4/X'_7$, $X'_5/X'_6$ represent carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound;

L represents an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof; and,

P represents zero to ten improved polyamides of claim 1.

11. A tandem-linked polyamide of claim 1 having the formula:



wherein γ is -NH-$CH_2$-$CH_2$-$CH_2$-CONH- hairpin linkage derived from γ-aminobutyric acid or a chiral hairpin linkage derived from R-2,4-diaminobutyric acid;

$X_1/X_{12}$, $X_2/X_{11}$, $X_3/X_{10}$, $X_4/X_9$, $X_5/X_8$, $X_6/X_7$, $X'_1/X'_{12}$, $X'_2/X'_{11}$, $X'_3/X'_{10}$, $X'_4/X'_9$, $X'_5/X'_8$ and $X'_6/X'_7$ represent carboxamide binding pairs which bind DNA base pairs wherein at least one binding pair is Hp/Py or Py/Hp and the other is selected

from the group consisting of Py/Im, Im/Py, and Py/Py to correspond to the DNA base pair in the minor groove to be bound;

L represents an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof; and,

P represents zero to ten improved polyamides of claim 1.

12. A tandem-linked polyamide comprising a first and second polyamide wherein said first polyamide is a polyamide of claim 4 and said second polyamide is selected from the group consisting of a polyamide of claim 5, 6 and 7; said first and second polyamides being linked by an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof bound to the γ-residue of said first polyamide and the carboxy tail of said second polyamide.

13. A tandem-linked polyamide comprising a first and second polyamide wherein said first polyamide is a polyamide of claim 5 and said second polyamide is selected from the group consisting of a polyamide of claim 4, 6 and 7; said first and second polyamides being linked by an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof bound to the γ-residue of said first polyamide and the carboxy tail of said second polyamide.

14. A tandem-linked polyamide comprising a first and second polyamide wherein said first polyamide is a polyamide of claim 6 and said second polyamide is selected from the group consisting of a polyamide of claim 4, 5, and 7; said first and second polyamides being linked by an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof bound to the γ-residue of said first polyamide and the carboxy tail of said second polyamide.

15. A tandem-linked polyamide comprising a first and second polyamide wherein said first polyamide is a polyamide of claim 7 and said second polyamide is selected from the group consisting of a polyamide of claim 4, 5, and 6; said first and second polyamides being linked by an amino acid linking group selected from the group consisting of β-alanine, 5-aminovaleric acid (δ) and a derivative thereof bound to the γ-residue of said first polyamide and the carboxy tail of said second polyamide.

16. A tandem-linked polyamide of claims 8, 9, 10 or 11 wherein P represents zero to eight polyamides of claim 1.

17. A tandem-linked polyamide of claims 8, 9, 10 or 11 wherein P represents zero to six polyamides of claim 1.

18. A tandem-linked polyamide of claims 8, 9, 10 or 11 wherein P represents zero to four polyamides of claim 1.

19. A tandem-linked polyamide of claims 8, 9, 10 or 11 wherein P represents zero to two polyamides of claim 1.

20. A polyamide of claim 1 wherein said chiral hairpin linkage derived from R-2,4-diaminobutyric acid.

21. A polyamide of claim 1 selected the group consisting of:

ImPyPy-γ-PyPyPy-β-Dp;

ImPyPy-(R)$^{H_2N}$γ-PyPyPy-β-Dp;

ImPyPy-(S)$^{H_2N}$γ-PyPyPy-β-Dp;

ImPyPy-(R)$^{H_2N}$γ-PyPyPy-β-EtOH;

ImPyPy-(R)$^{Ac}$γ-PyPyPy-β-Dp;

ImPyPy-(S)$^{Ac}$γ-PyPyPy-β-Dp;

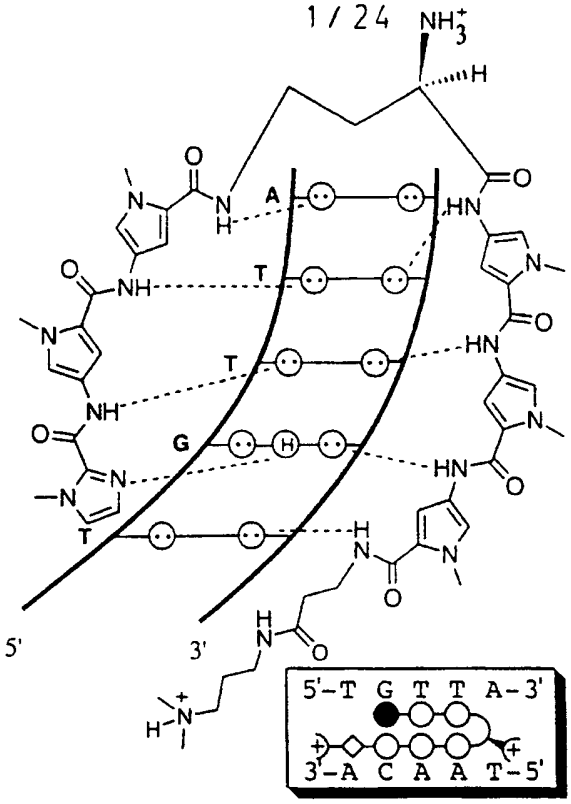
ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-β-]$^{HN}$γ-PyPyPy-β-Dp;
ImPyPy-(*R*)[ImPyPy-(R)$^{H2N}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp;
ImPyPy-(*R*)[ImPyPy-(R)$^{EDTA}$γ–PyPyPy-δ-]$^{HN}$γ-PyPyPy-β-Dp; and,
the pharmacologically acceptable salts thereof.

22. A polyamide of claim 1 further comprising an R-2 amino group attached to a detectable label.

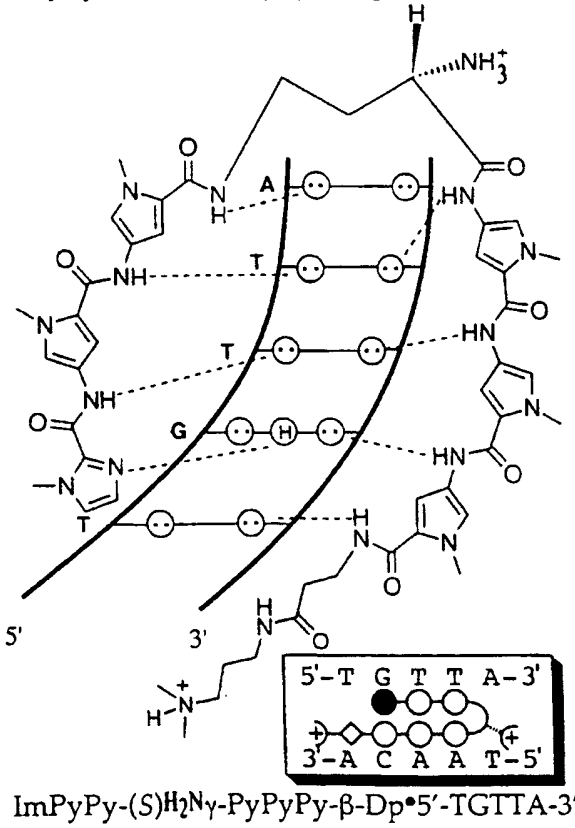
23. A method of increasing the affinity polyamide having a hairpin turn for binding a DNA sequence comprising substituting the α-position of the γ-aminobutyric acid with (R)-2,4-diaminobutyric acid

24. A method of claim 8 wherein the α-position of the γ-aminobutyric acid residue of the hairpin is substituted with an acetylated R-enantiomer of aminobutyric acid.

25. A method of inhibiting gene expression comprising contacting a regulatory sequence of a gene with a polyamide of claim 1.




ImPyPy-(*R*)$H_2N\gamma$-PyPyPy-β-Dp•5′-TGTTA-3′ FIG. 1A




ImPyPy-(*S*)$H_2N\gamma$-PyPyPy-β-Dp•5′-TGTTA-3′ FIG. 1B

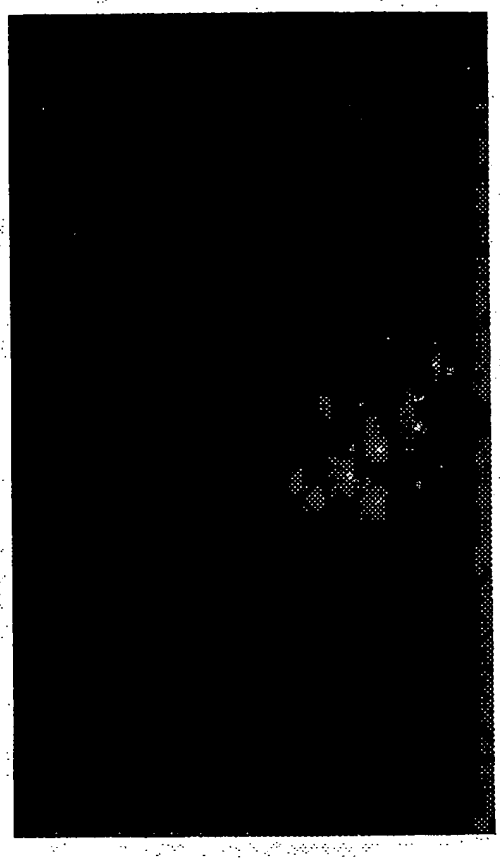


FIG. 2A

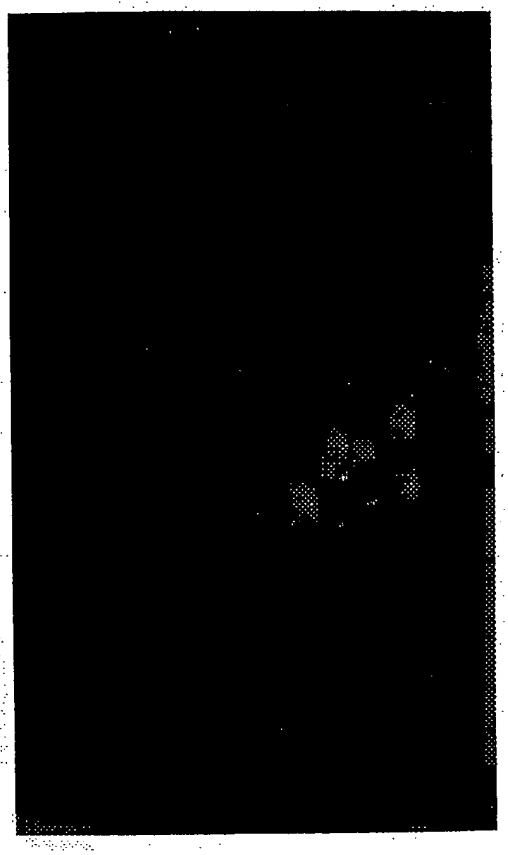
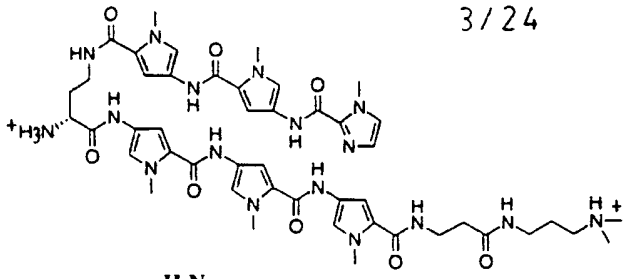
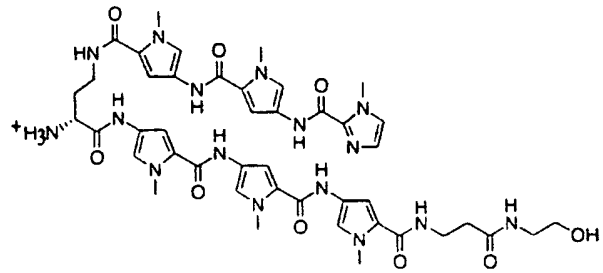


FIG. 2B



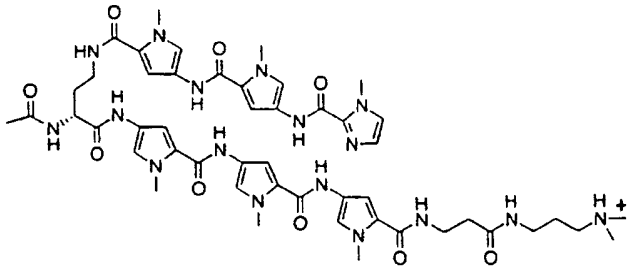
ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Dp (1-*R*)

FIG. 3A



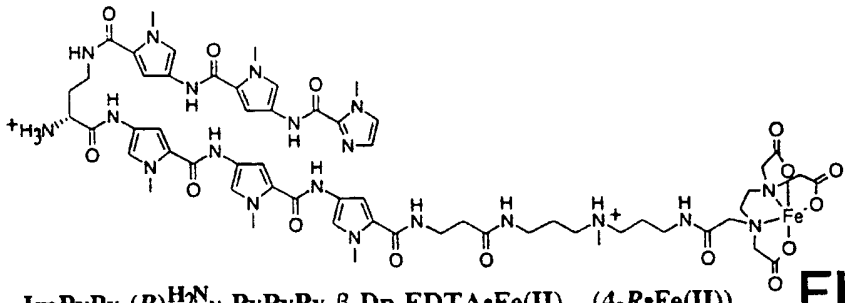
ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-EtOH (2-*R*)

FIG. 3B



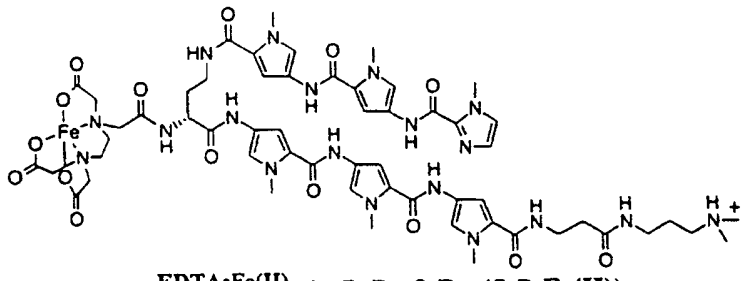
ImPyPy-(*R*)$^{Ac}$γ-PyPyPy-β-Dp (3-*R*)

FIG. 3C



ImPyPy-(*R*)$^{H_2N}$γ-PyPyPy-β-Dp-EDTA•Fe(II) (4-*R*•Fe(II))

FIG. 3D



ImPyPy-(*R*)$^{EDTA•Fe(II)}$γ-PyPyPy-β-Dp (5-*R*•Fe(II))

FIG. 3E

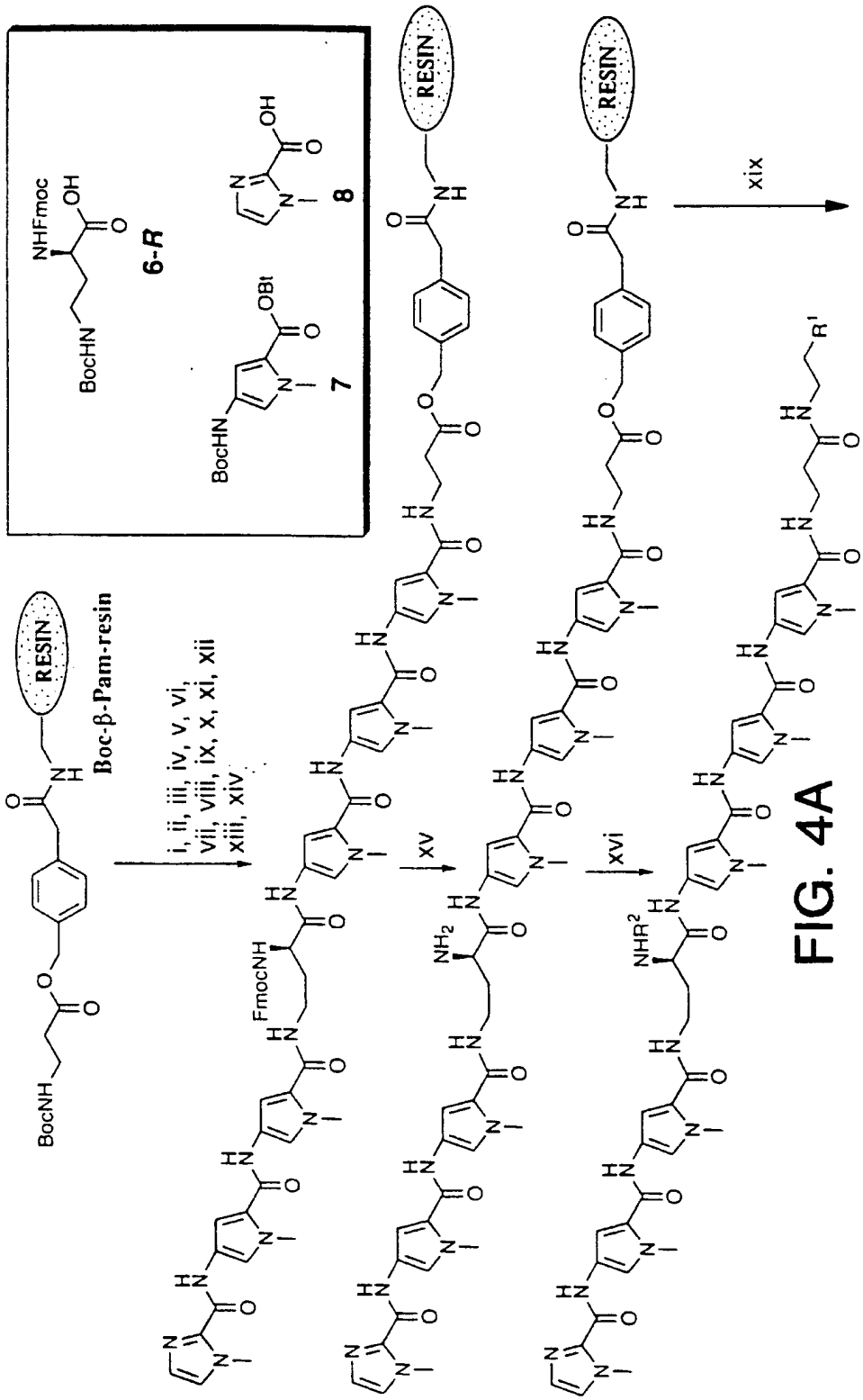

RESIN
Boc-β-Pam-resin
BocNH
i, ii, iii, iv, v, vi,
vii, viii, ix, x, xi, xii
xiii, xiv
FmocNH
xv
NH2
xvi
NHR2
xix
R'
6-R
NHFmoc
OH
BocHN
8
OBt
7

FIG. 4A

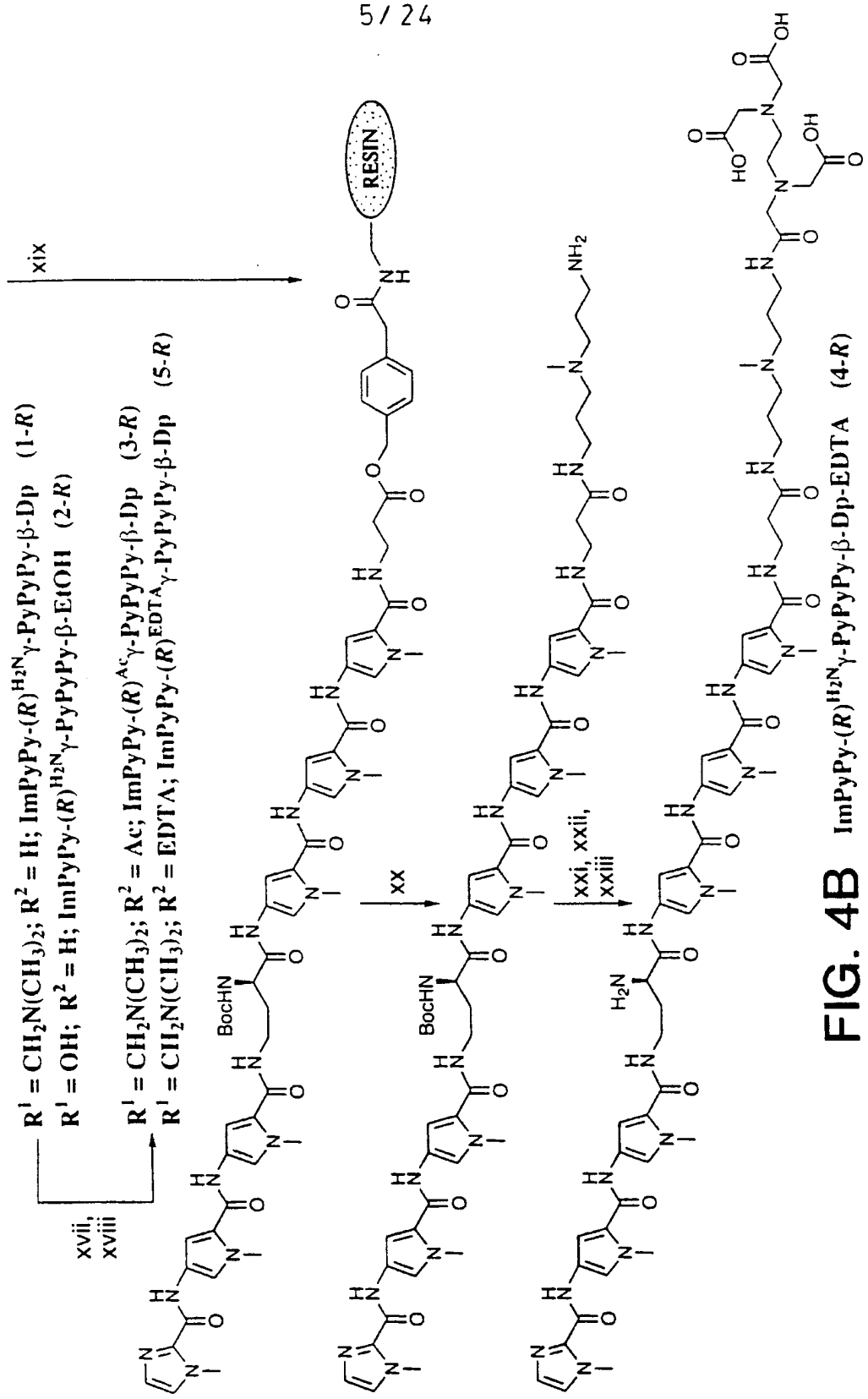



FIG. 4B ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA (4-R)

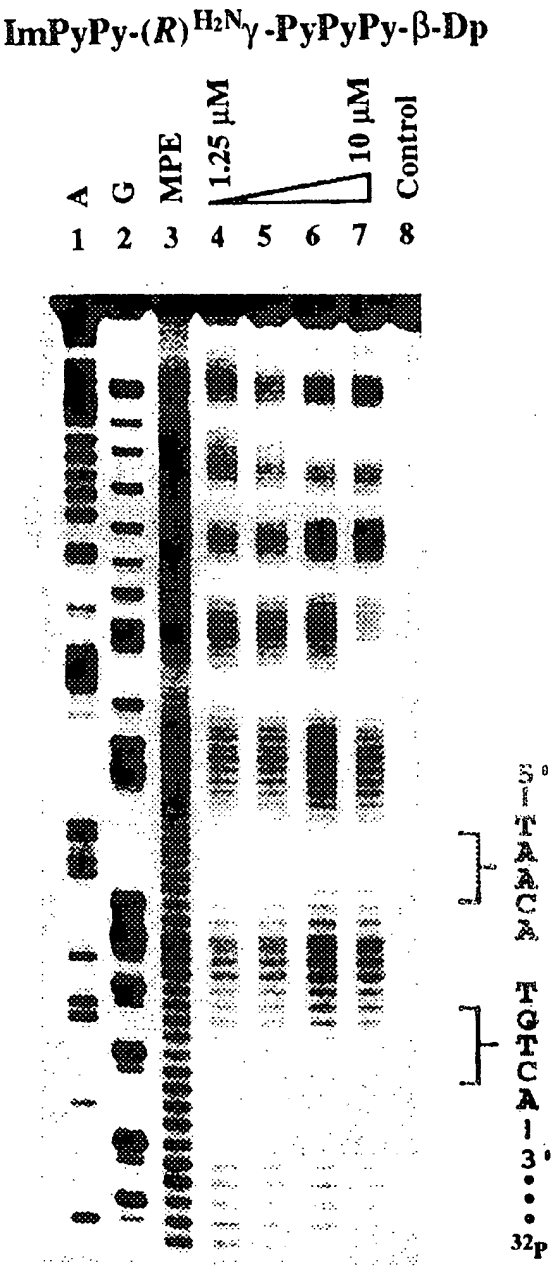

ImPyPy-(R)H2Nγ-PyPyPy-β-Dp
A
G
MPE
1.25 μM
10 μM
Control
1 2 3 4 5 6 7 8
5'-TAACA TGTCA-3'
32P

FIG. 5A

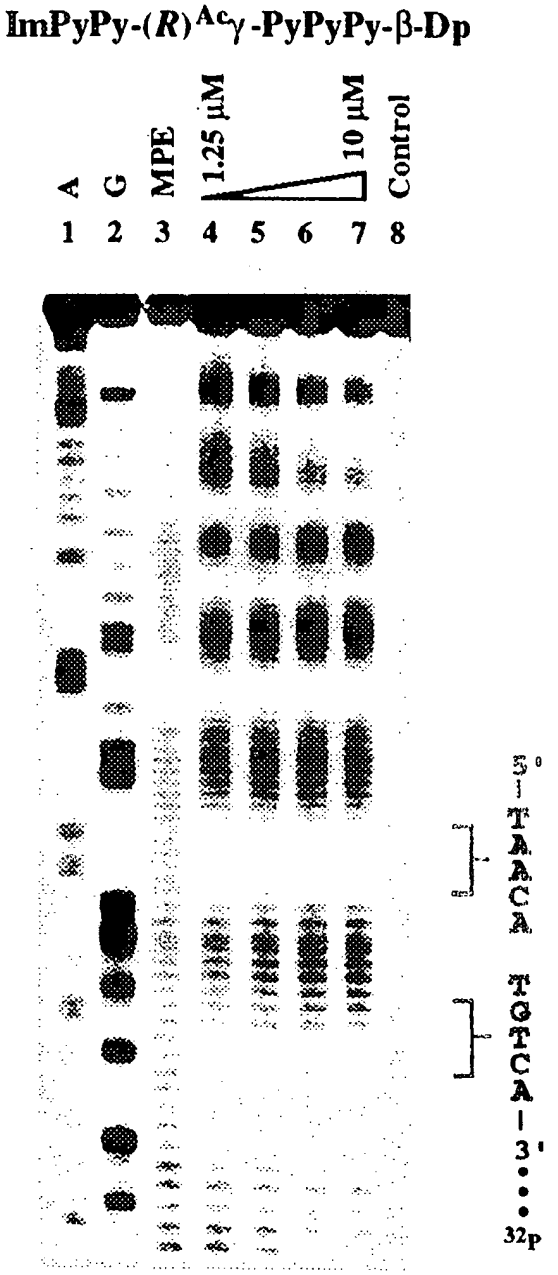

ImPyPy-(R)Acγ-PyPyPy-β-Dp
A
G
MPE
1.25 μM
10 μM
Control
1 2 3 4 5 6 7 8
5'-TAACA TGTCA-3'
32P

FIG. 5B

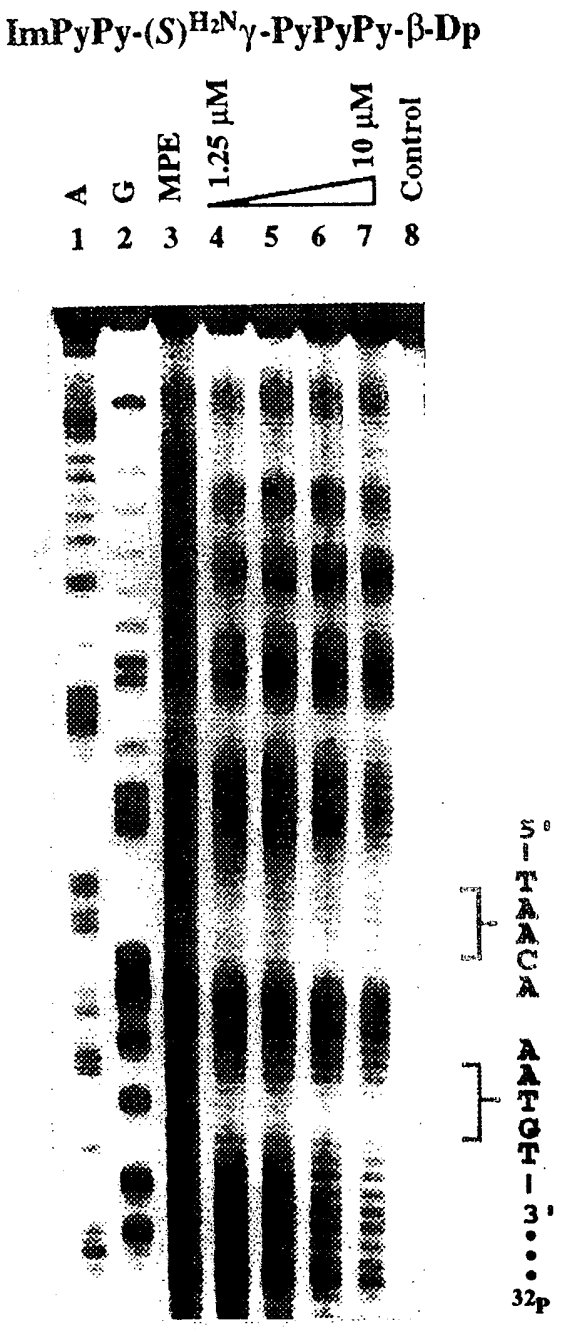

ImPyPy-(S)H2Nγ-PyPyPy-β-Dp
A G MPE 1.25 μM 10 μM Control
1 2 3 4 5 6 7 8
5'-TAACA AATGT-3'
32P

FIG. 5C

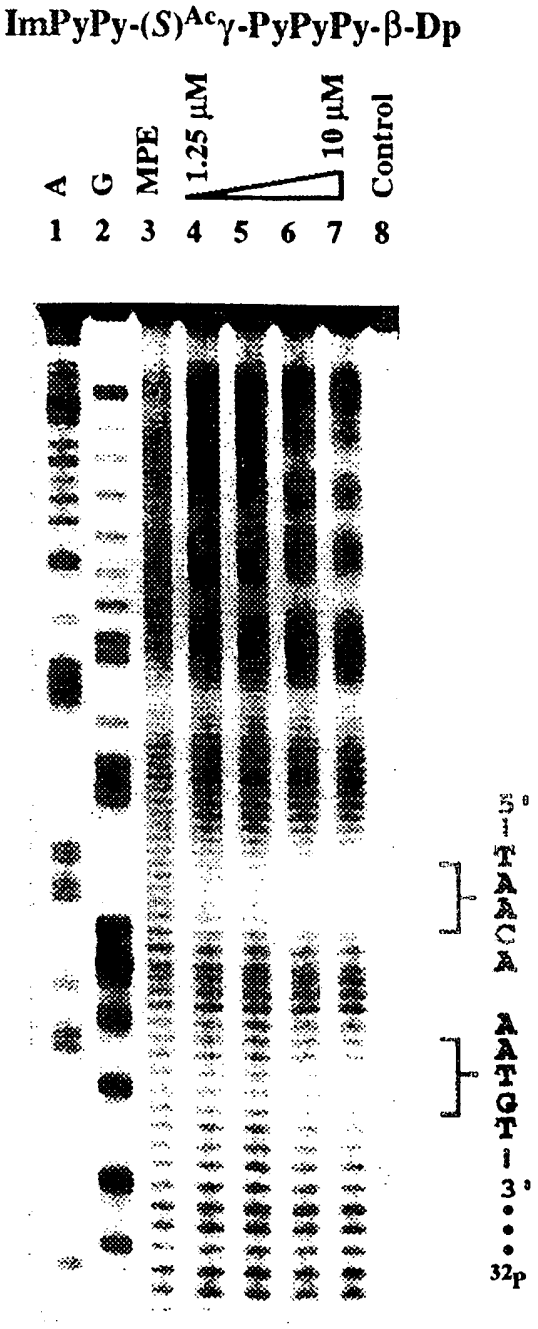

ImPyPy-(S)Acγ-PyPyPy-β-Dp
A G MPE 1.25 μM 10 μM Control
1 2 3 4 5 6 7 8
5'-TAACA AATGT-3'
32P

FIG. 5D

EcoRI
135 bp
BsrBI
32P
ImPyPy-(R)H2Nγ-PyPyPy-β-Dp (1.25μM)
32P 5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'
ImPyPy-(R)Acγ-PyPyPy-β-Dp (1.25μM)
32P 5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'
ImPyPy-(S)H2Nγ-PyPyPy-β-Dp (1.25μM)
32P 5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'
ImPyPy-(S)Acγ-PyPyPy-β-Dp (2.5μM)
32P 5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'

FIG. 6

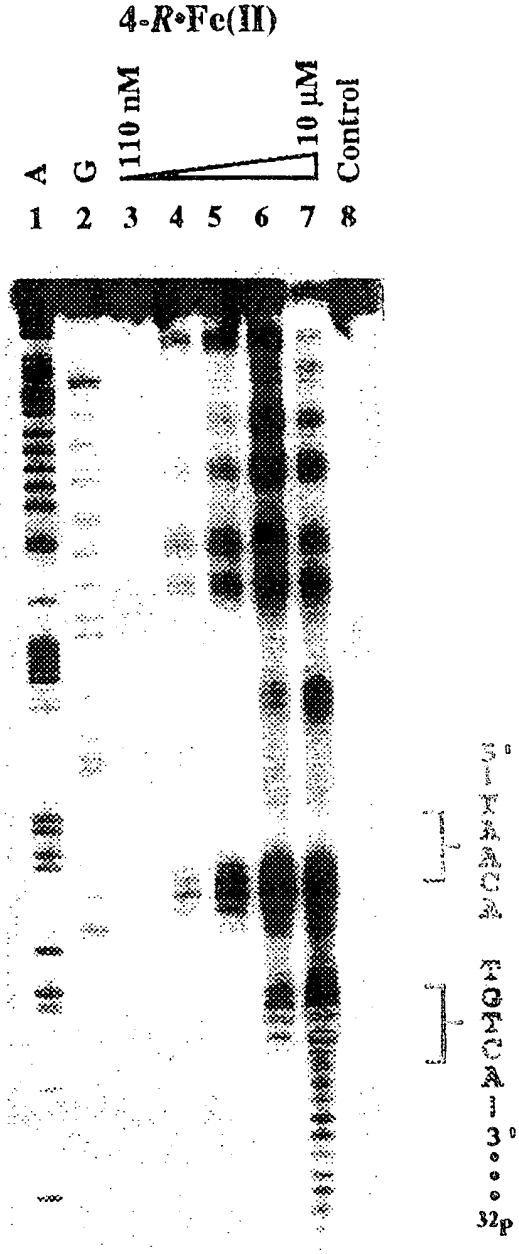

4-R•Fe(II)
A G 110 nM 10 μM Control
1 2 3 4 5 6 7 8
5'-TAGGCA TGTCCA-3'
32P

FIG. 7A

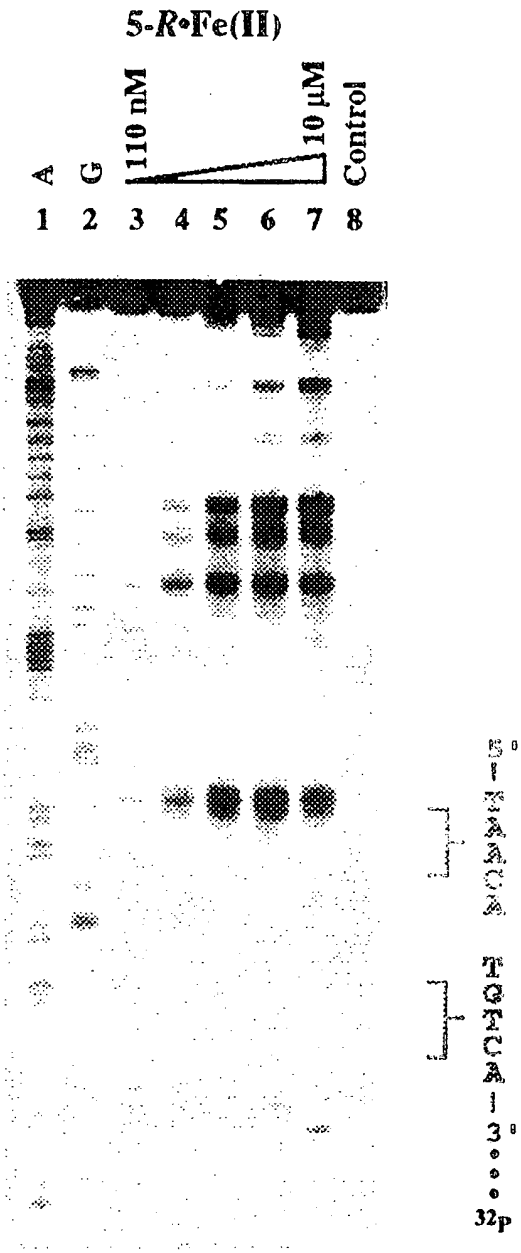

5-R•Fe(II)
A G 110 nM 10 μM Control
1 2 3 4 5 6 7 8
5'-TAGGCA TGTCCA-3'
32P

FIG. 7B

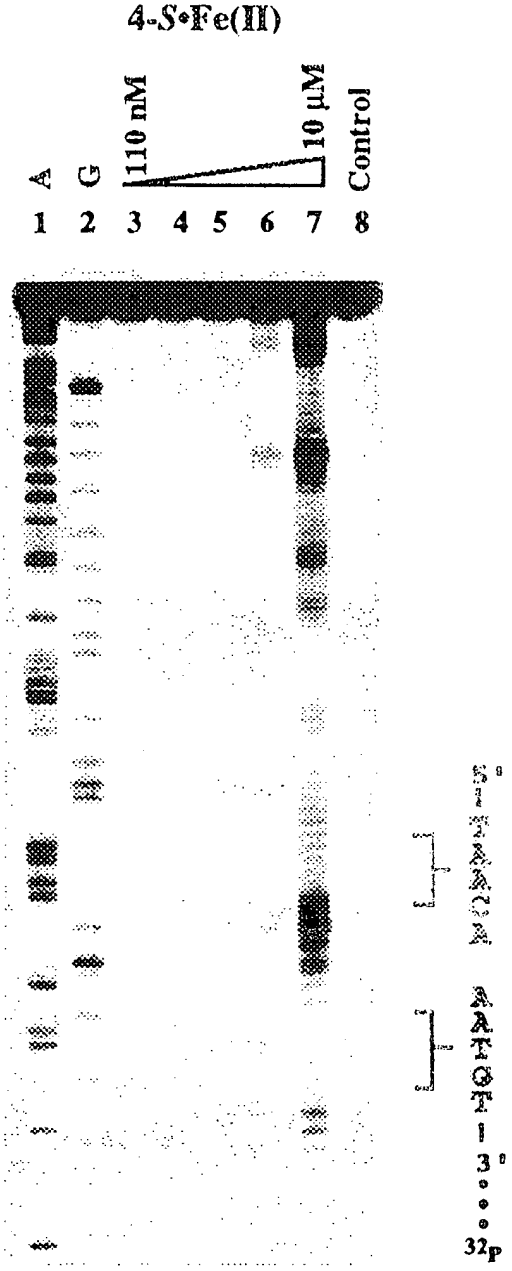

4-S•Fe(II)
A G 110 nM 10 μM Control
1 2 3 4 5 6 7 8
5'-Endlabel
AFHOF – 3'
32P

FIG. 7C

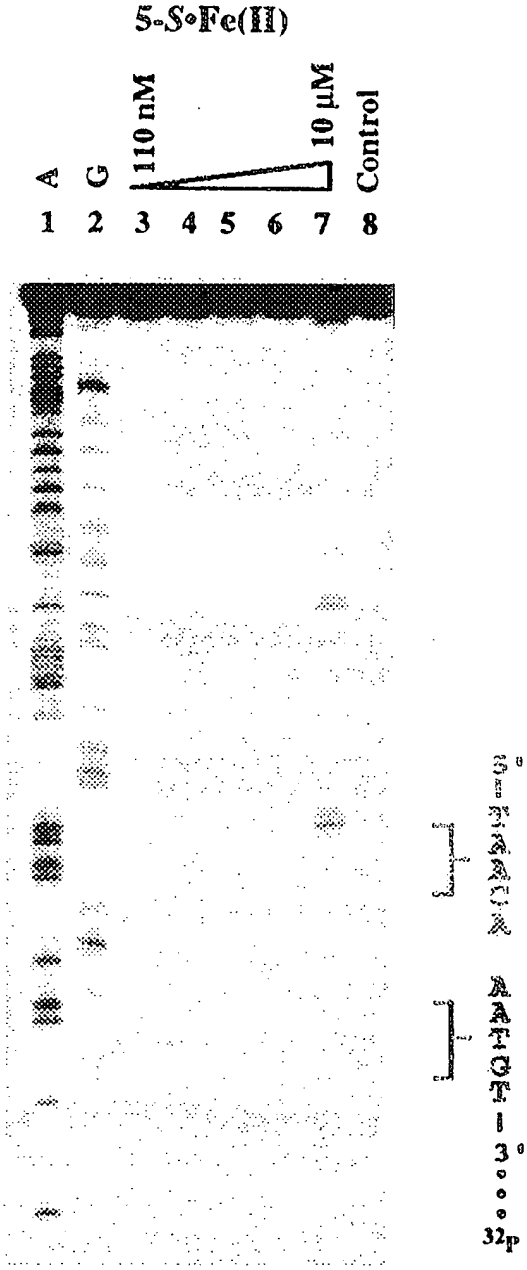

5-S•Fe(II)
A G 110 nM 10 μM Control
1 2 3 4 5 6 7 8
5'-Endlabel
AFHOF – 3'
32P

FIG. 7D

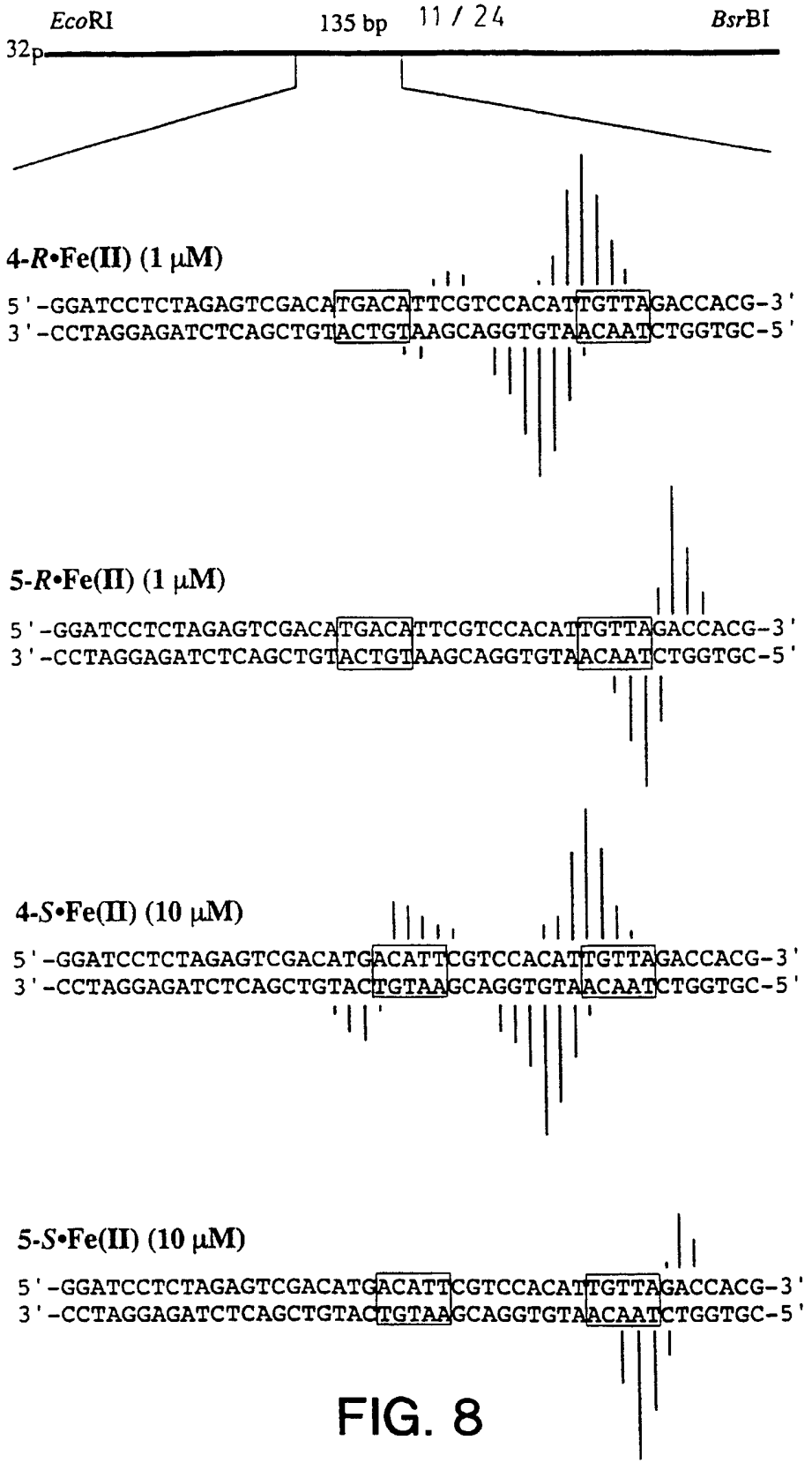
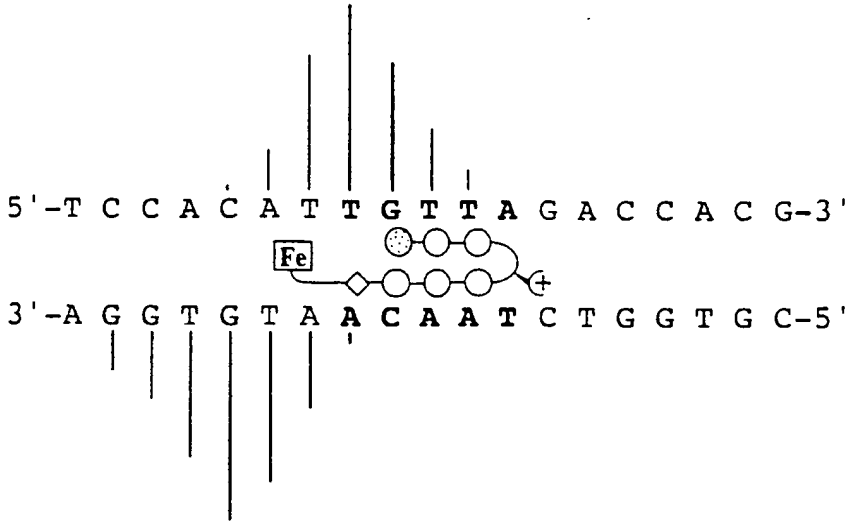
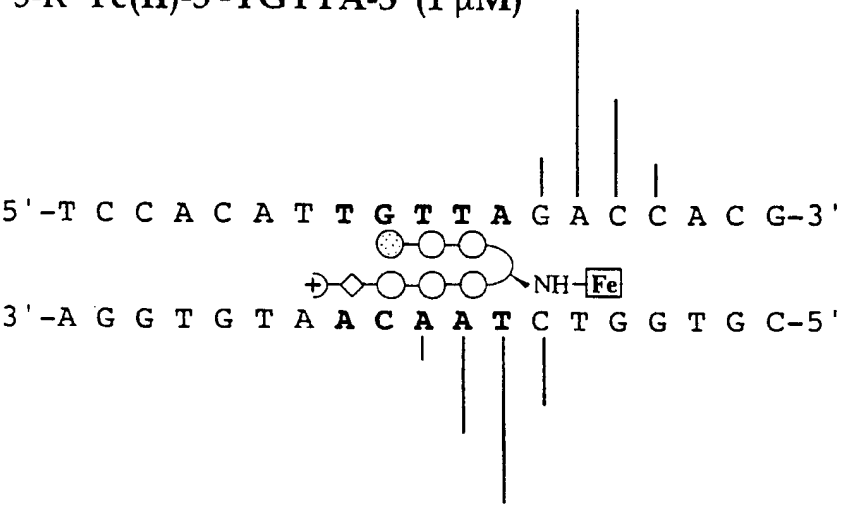

EcoRI
135 bp
BsrBI
32p
4-R•Fe(II) (1 μM)
5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'
5-R•Fe(II) (1 μM)
5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'
4-S•Fe(II) (10 μM)
5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'
5-S•Fe(II) (10 μM)
5'-GGATCCTCTAGAGTCGACATGACATTCGTCCACATTGTTAGACCACG-3'
3'-CCTAGGAGATCTCAGCTGTACTGTAAGCAGGTGTAACAATCTGGTGC-5'

FIG. 8

4-*R*•Fe(II)-5'-TGTTA-3' (1 μM)




5-*R*•Fe(II)-5'-TGTTA-3' (1 μM)




4-*S*•Fe(II)-5'-ACATT-3' (10 μM)

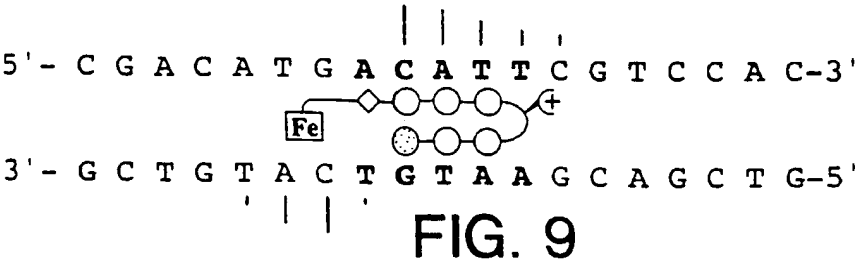



FIG. 9

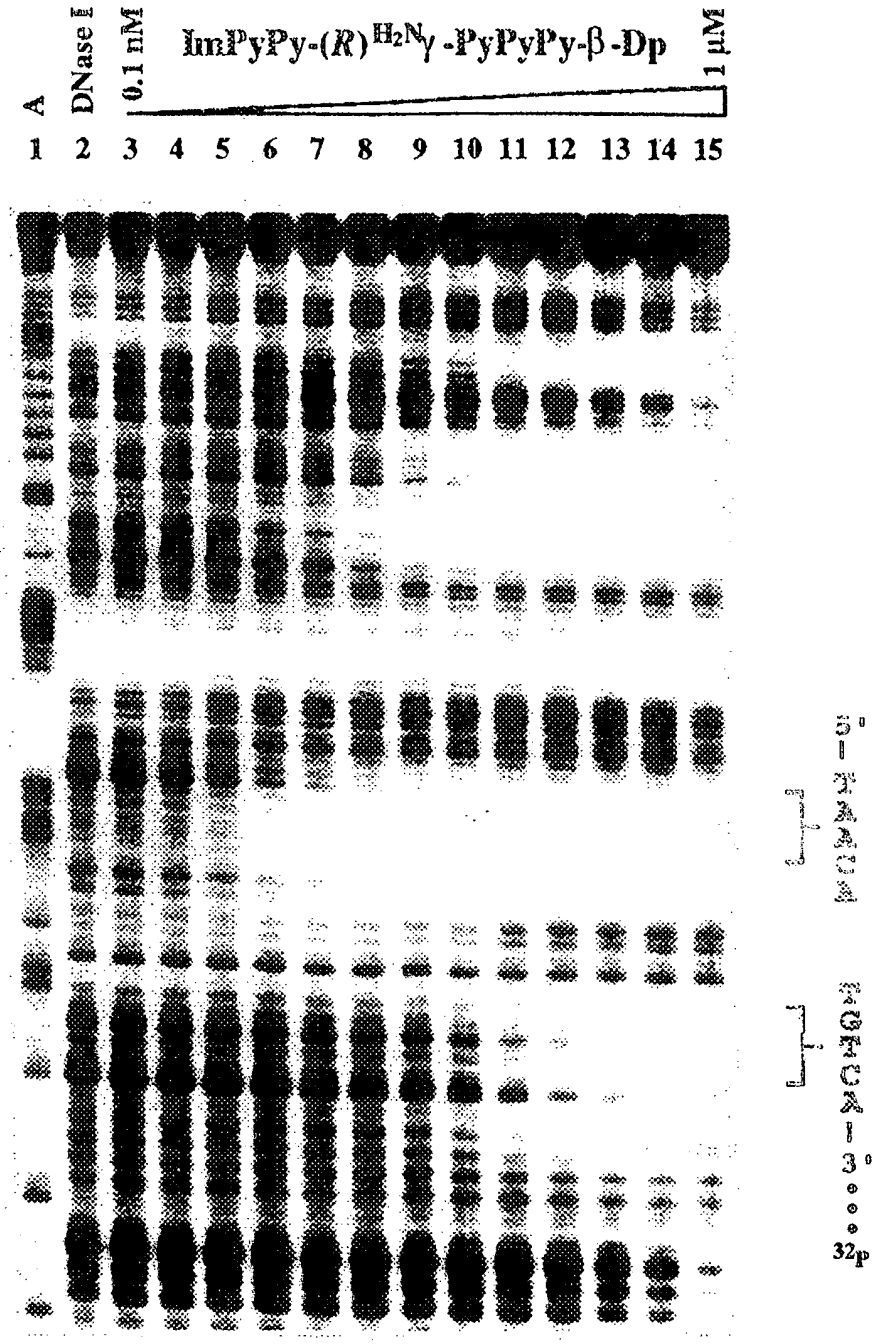

A
DNase I
0.1 nM
ImPyPyPy-(R)-H2Nγ-PyPyPyPy-β-Dp
1 μM
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
5'-AAGCTT
3'-TTCGAA
32P

FIG. 10A

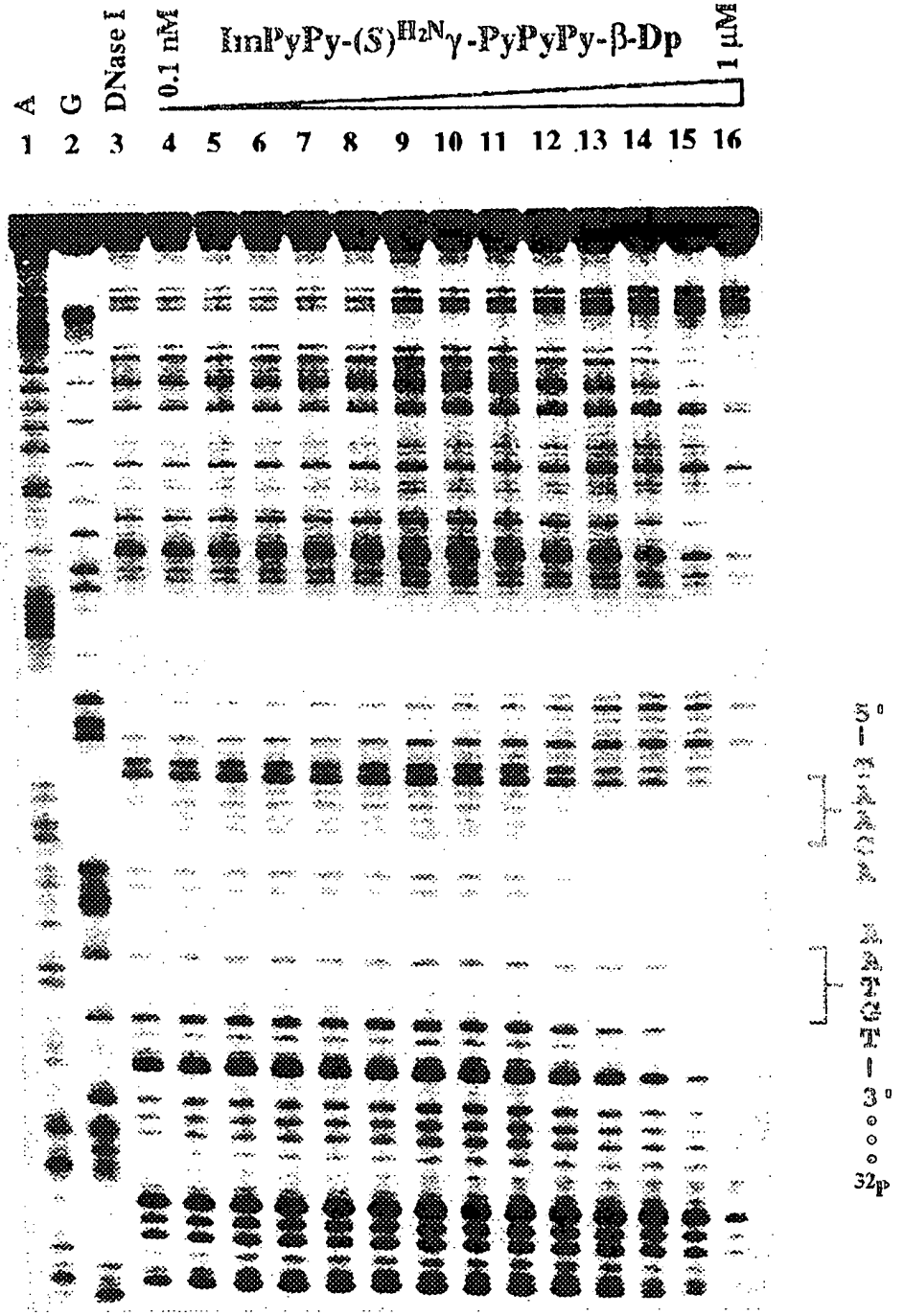

A
G
DNase I
0.1 nM
ImPyPy-(S)H2Nγ-PyPyPy-β-Dp
1 μM
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
5'
3'
32p

FIG. 10B

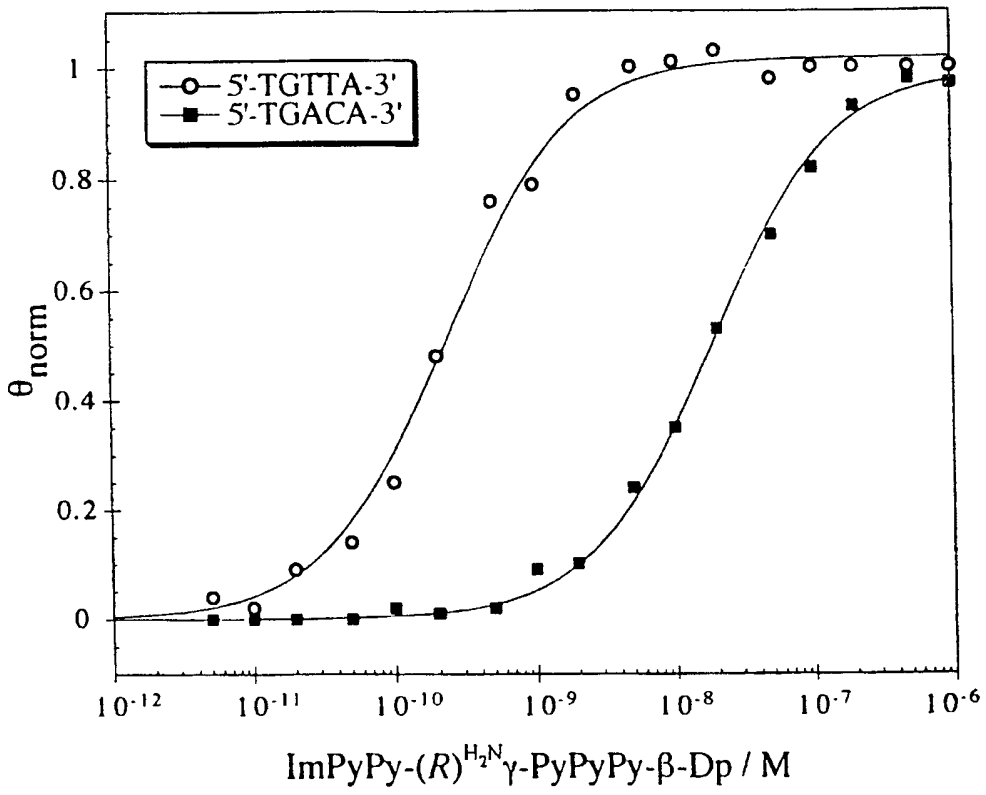

5'-TGTTA-3'
5'-TGACA-3'
θnorm
1
0.8
0.6
0.4
0.2
0
10-12
10-11
10-10
10-9
10-8
10-7
10-6
ImPyPy-(R)H2Nγ-PyPyPy-β-Dp / M

FIG. 11

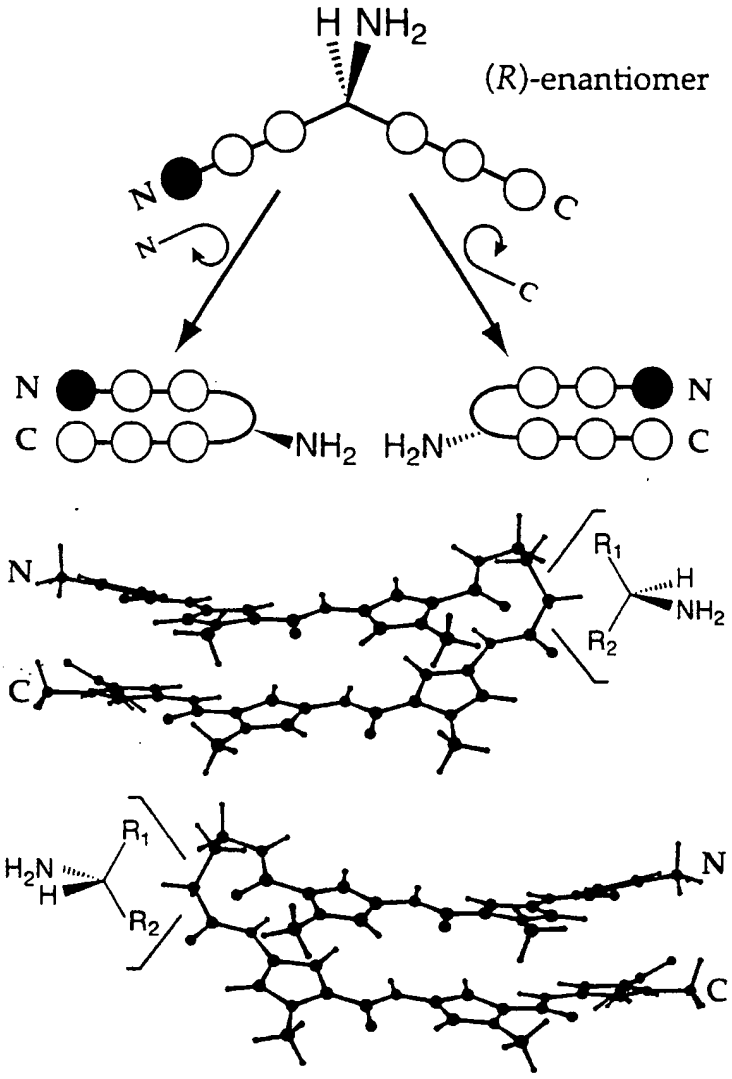



FIG. 12

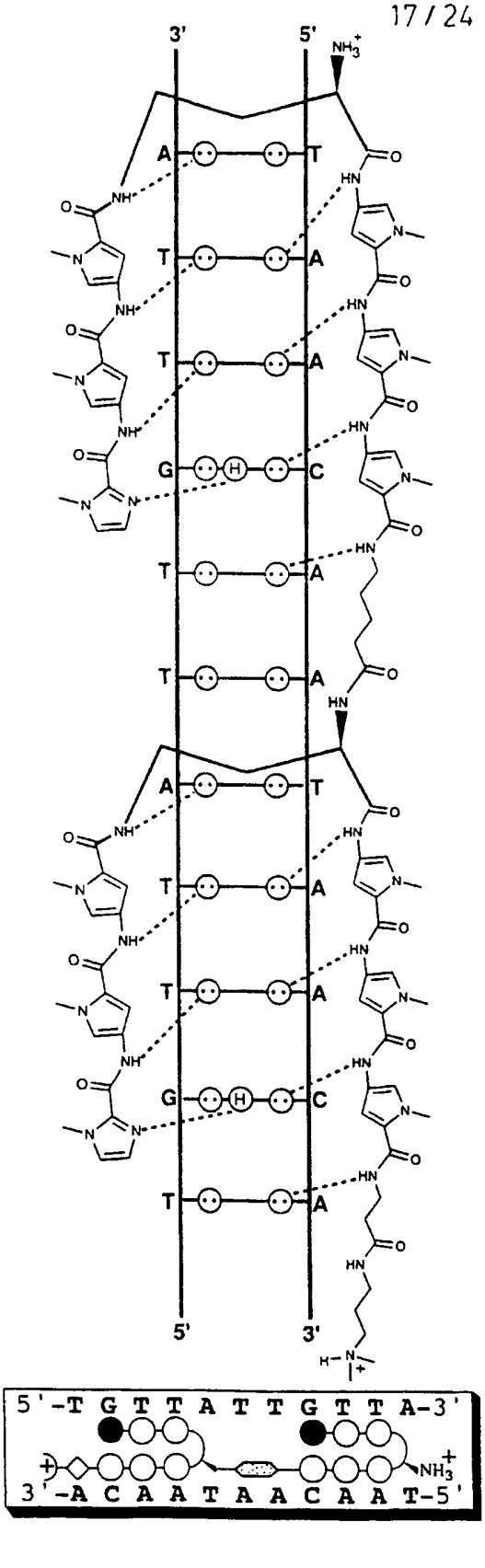
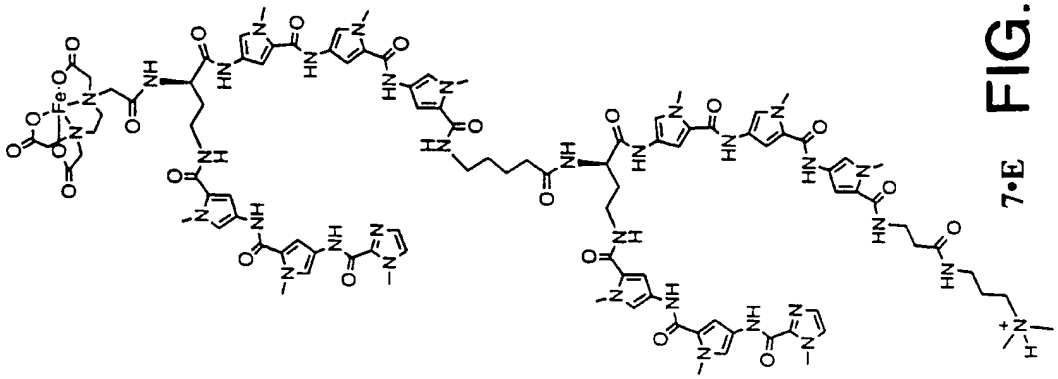
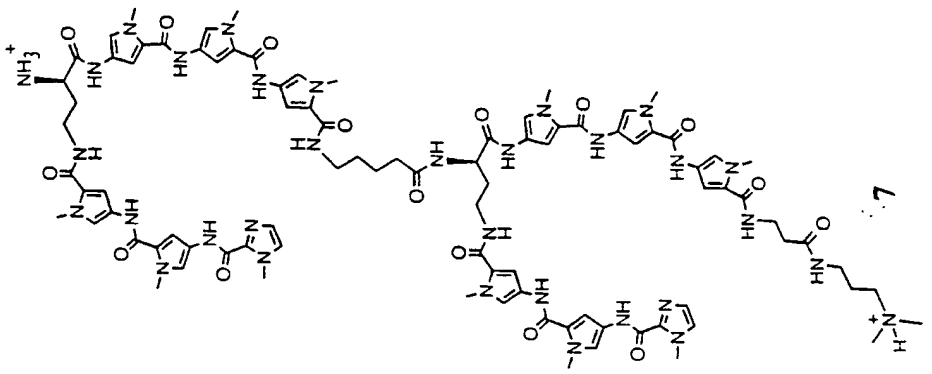


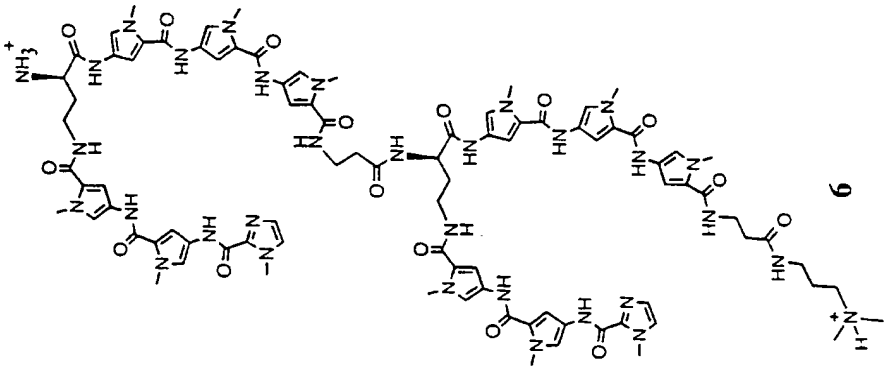

FIG. 13



7•E



7



6

FIG. 14

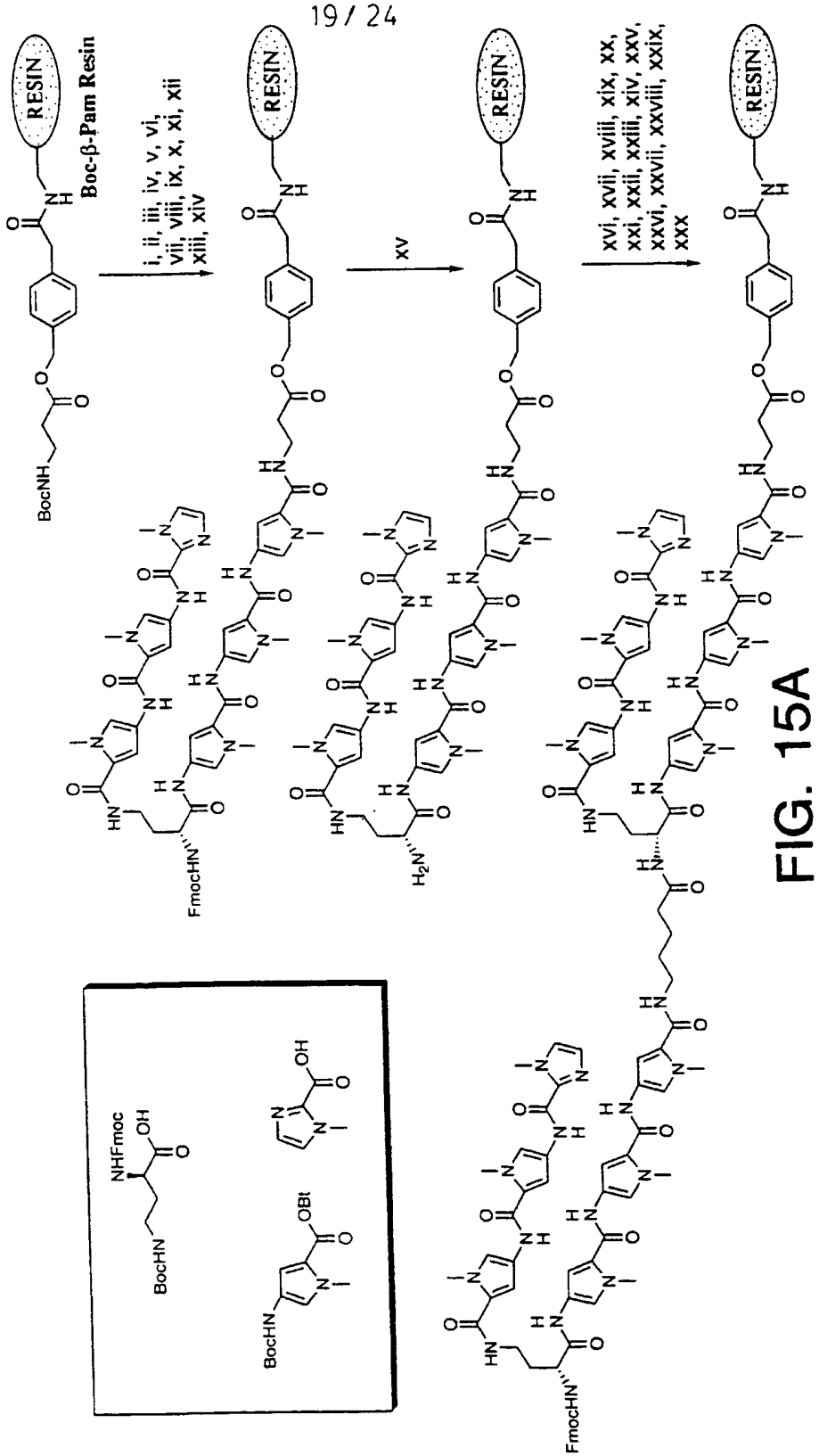

RESIN
Boc-β-Pam Resin
i, ii, iii, iv, v, vi,
vii, viii, ix, x, xi, xii
xiii, xiv
xv
xvi, xvii, xviii, xix, xx,
xxi, xxii, xxiii, xiv, xxv,
xxvi, xxvii, xxviii, xxix,
xxx
BocNH
FmocHN
H2N
NHFmoc
BocHN
OBt

FIG. 15A

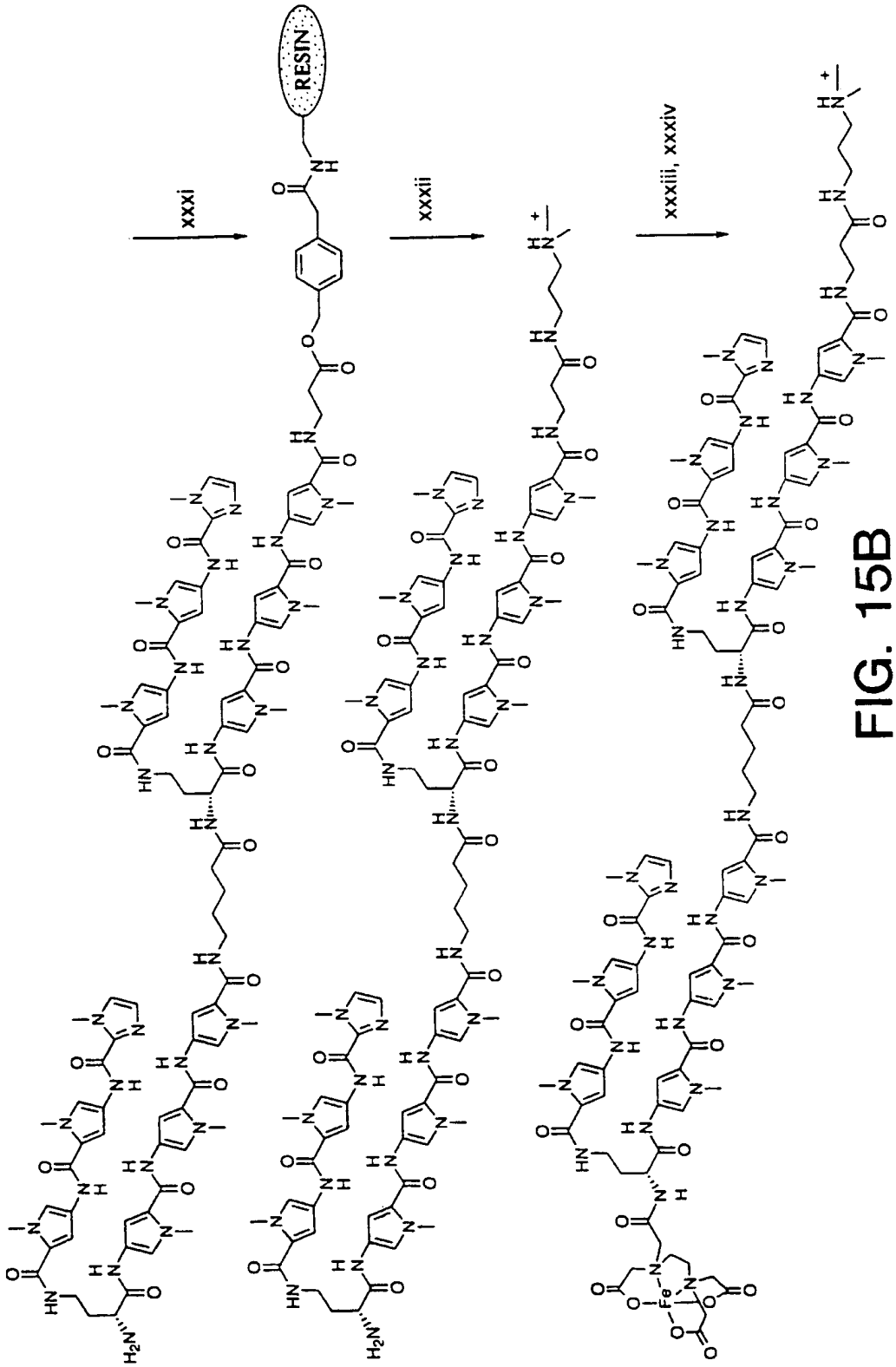



FIG. 15B

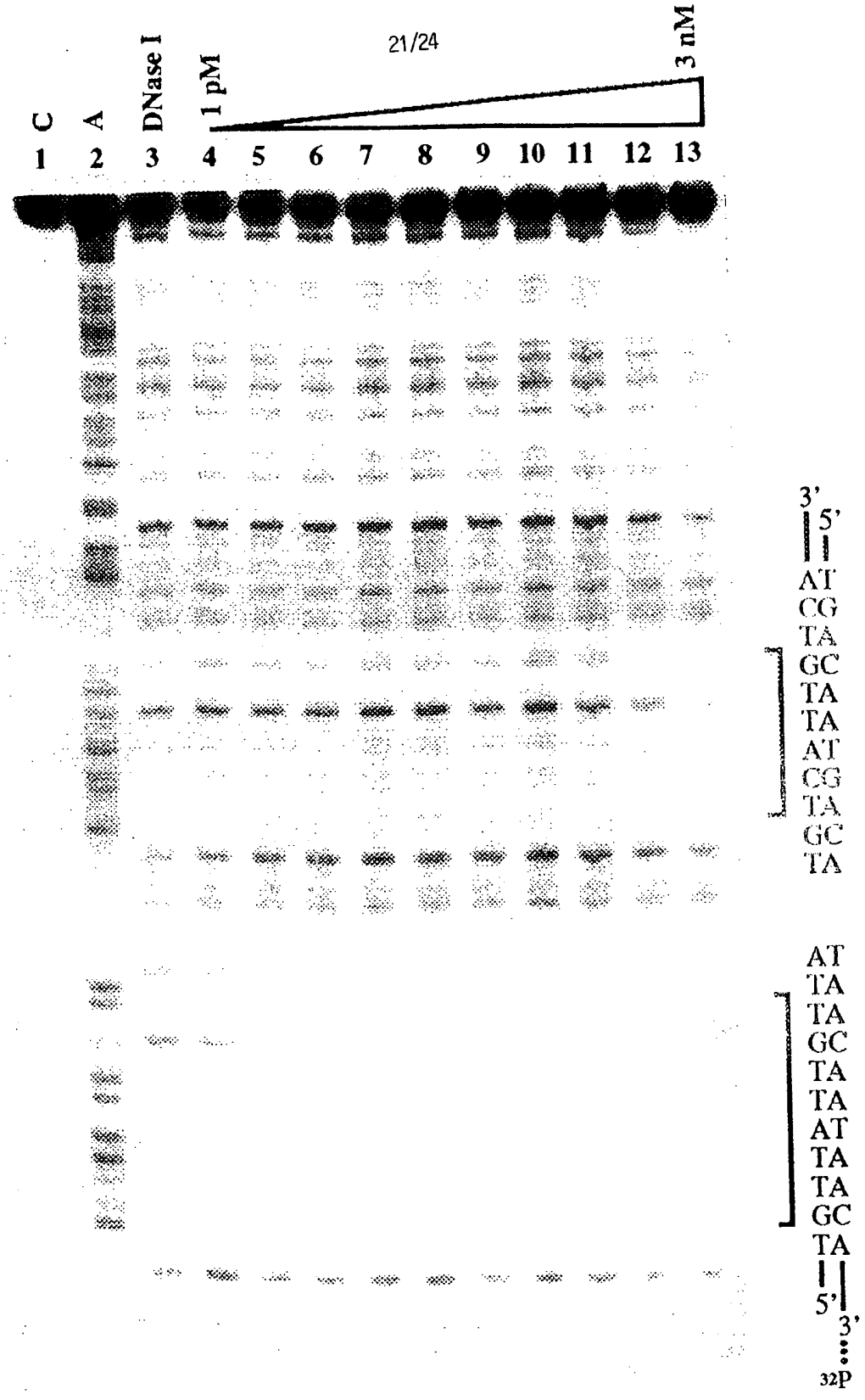

C
A
DNase I
1 pM
3 nM
1 2 3 4 5 6 7 8 9 10 11 12 13
3'
5'
AT
CG
TA
GC
TA
TA
AT
CG
TA
GC
TA
AT
TA
TA
GC
TA
TA
AT
TA
TA
GC
TA
5'
3'
32P

FIG. 16

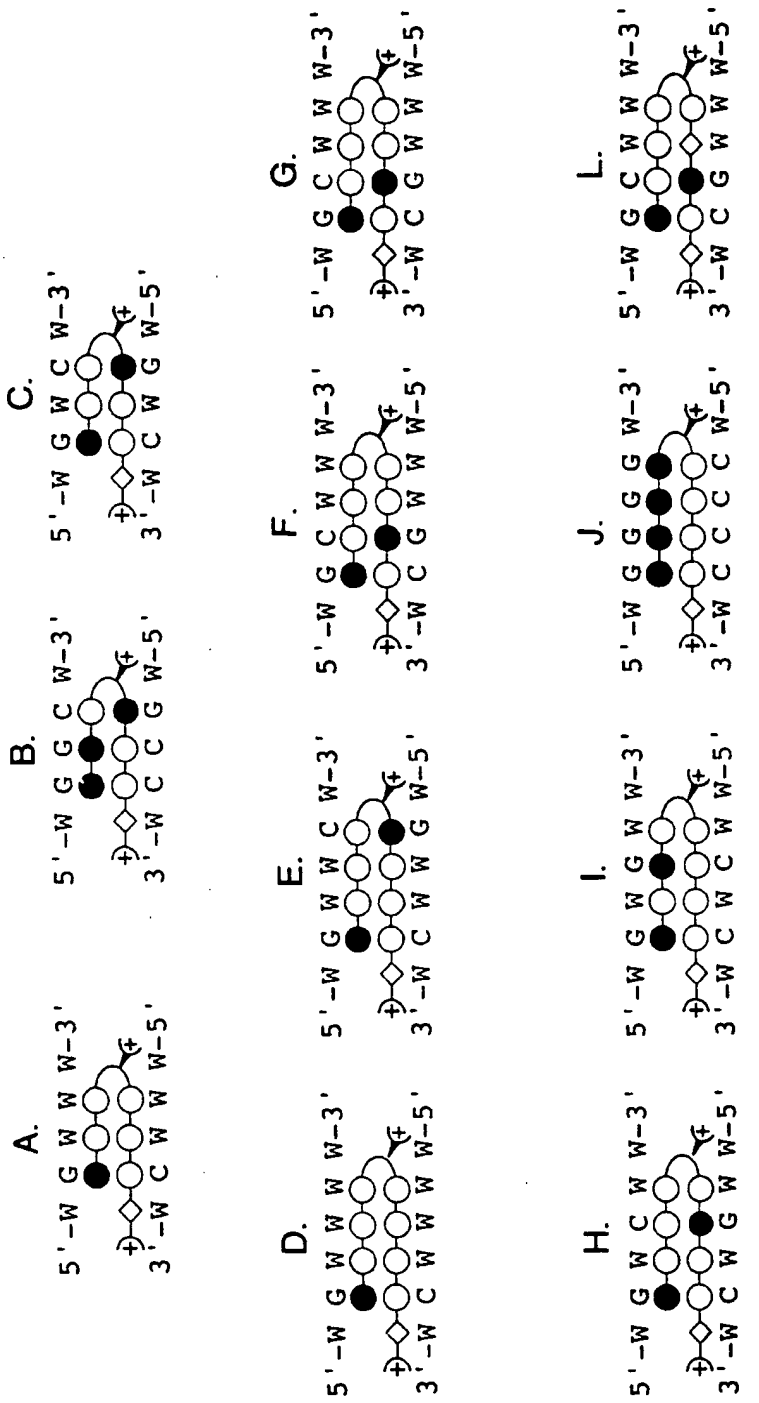

A.
5'-W G W W W-3'
3'-W C W W W-5'
B.
5'-W G G C W-3'
3'-W C C G W-5'
C.
5'-W G W C W-3'
3'-W C W G W-5'
D.
5'-W G W W W-3'
3'-W C W W W-5'
E.
5'-W G W W C W-3'
3'-W C W W G W-5'
F.
5'-W G C W W W-3'
3'-W C G W W W-5'
G.
5'-W G C W W W-3'
3'-W C G W W W-5'
H.
5'-W G W C W W-3'
3'-W C W G W W-5'
I.
5'-W G W G W W-3'
3'-W C W C W W-5'
J.
5'-W G G G G W-3'
3'-W C C C C W-5'
L.
5'-W G C W W W-3'
3'-W C G W W W-5'

FIG. 17A

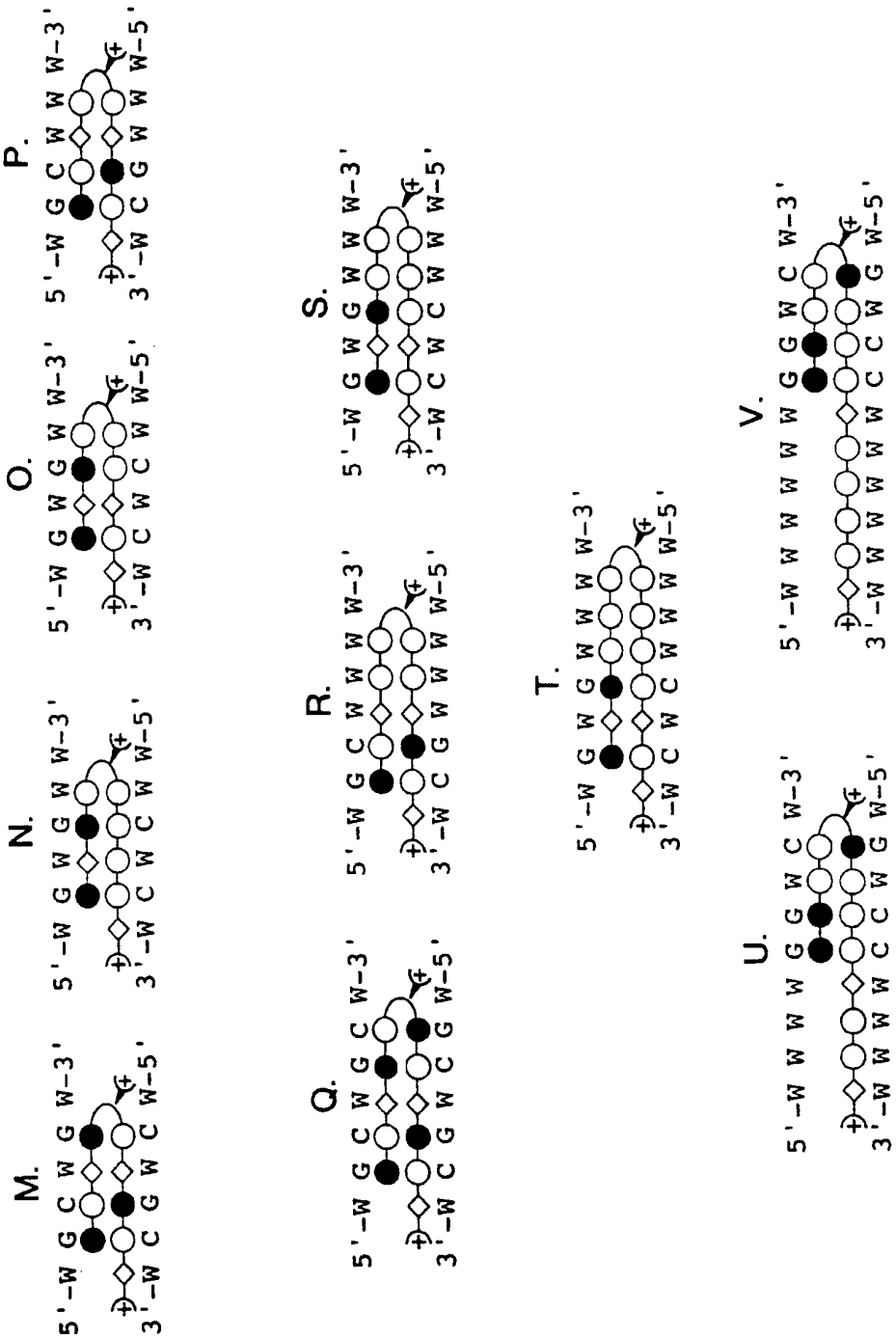

M.
5'-W G C W G W-3'
3'-W C G W C W-5'
N.
5'-W G W G W G W W-3'
3'-W C W C W C W W-5'
O.
5'-W G W G W G W W-3'
3'-W C W C W C W W-5'
P.
5'-W G C W W W-3'
3'-W C G W W W-5'
Q.
5'-W G C W G C W-3'
3'-W C G W C G W-5'
R.
5'-W G C W W W W-3'
3'-W C G W W W W-5'
S.
5'-W G W G W W W-3'
3'-W C W C W W W-5'
T.
5'-W G W G W W W W-3'
3'-W C W C W W W W-5'
U.
5'-W W W W G G W C W-3'
3'-W W W W C C W G W-5'
V.
5'-W W W W W G G W C W-3'
3'-W W W W W C C W G W-5'

FIG. 17B

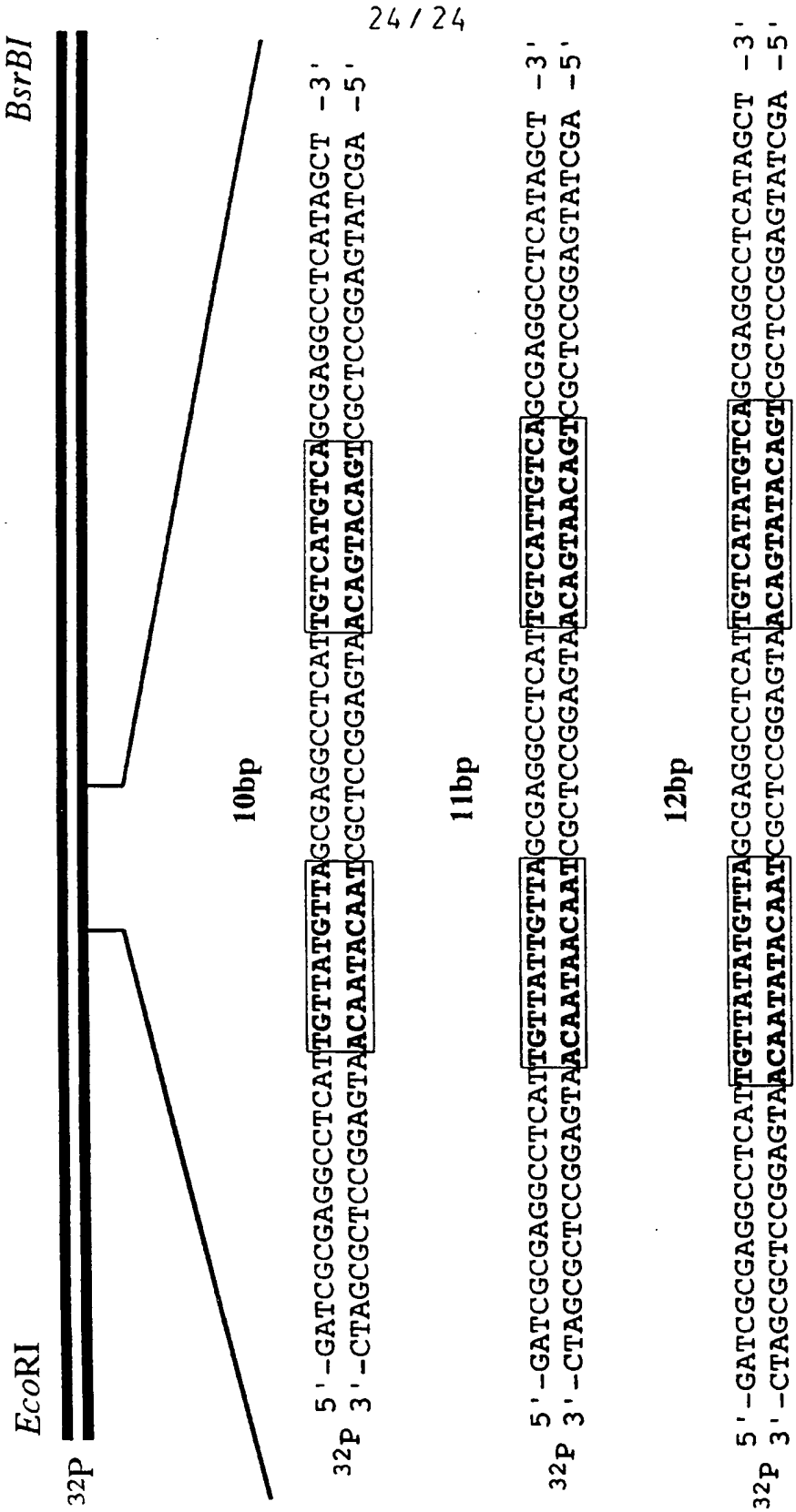

BsrBI
EcoRI
32P
10bp
5'-GATCGCGAGGCCTCATTGTTATGTTAGCGAGGCCTCATTGTCATGTCAGCGAGGCCTCATAGCT -3'
32P 3'-CTAGCGCTCCGGAGTAACAATACAATCGCTCCGGAGTAACAGTACAGTCGCTCCGGAGTATCGA -5'
11bp
5'-GATCGCGAGGCCTCATTGTTATTGTTAGCGAGGCCTCATTGTCATTGTCAGCGAGGCCTCATAGCT -3'
32P 3'-CTAGCGCTCCGGAGTAACAATAACAATCGCTCCGGAGTAACAGTAACAGTCGCTCCGGAGTATCGA -5'
12bp
5'-GATCGCGAGGCCTCATTGTTATATGTTAGCGAGGCCTCATTGTCATATGTCAGCGAGGCCTCATAGCT -3'
32P 3'-CTAGCGCTCCGGAGTAACAATATACAATCGCTCCGGAGTAACAGTATACAGTCGCTCCGGAGTATCGA -5'

FIG. 18

# INTERNATIONAL SEARCH REPORT

| | |
|---|---|
| | International Application No PCT/US 98/03829 |

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07D403/14 C07D207/34 A61K31/415 C12Q1/68

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07D A61K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| T | D. M. HERMAN ET AL : "Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 120, no. 7, 25 February 1998, DC US, pages 1382-1391, XP002071829 see page 1383, right-hand column - page 1385 -/-- | 1-25 |

[X] Further documents are listed in the continuation of box C.

[ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the International search report |
|---|---|
| 17 July 1998 | 09/09/1998 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Voyiazoglou, D |

3

Form PCT/ISA/210 (second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Ir. ational Application No
PCT/US 98/03829

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication,where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | S. E. SWALLEY ET AL : "Recognition of a 5'-(A,T)GGG(A,T)2-3' sequence in the minor groove of DNA by an eight-ring hairpin polyamide" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, vol. 118, no. 35, 4 September 1996, DC US, pages 8198-8206, XP002071830 cited in the application see abstract | 1-25 |

3

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

..ternational application No.
PCT/US 98/03829

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 23-25
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

**This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record**

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

**IMAGES ARE BEST AVAILABLE COPY.**
**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)